BUS. ADM. LIBRARY

JAN 7 1949

Journal of the

AMERICAN STATISTICAL ASSOCIATION

DECEMBER 1948

Commercial Uses of Sampling — *J. Stevens Stock and Joseph R. Hochstim* — 509

Variation of the Frequency of Fatal Quarrels with Magnitude — *Lewis F. Richardson* — 523

Bank Reserves and Business Fluctuations — *Clark Warburton* — 547

The Ordering of n Items Assigned to k Rank Categories by Votes of m Individuals — *Garret L. Schuyler* — 559

Levels of Significance for Variance Ratio of Two Samples of Equal Size — *C. J. Kirchen* — 564

Main Effects and Interactions — *D. J. Finney* — 566

A Test for Symmetry in Contingency Tables — *Albert H. Bowker* — 572

The War Production Board's Statistical Reporting Experience, Part IV — *David Novick and George A. Steiner* — 575

Correction to "On Estimating Precision of Measuring Instruments and Product Variability" — *Frank E. Grubbs* — 564

Statistical Methodology Index — *Oscar Krisen Buros* — 616

Index to Volume 43, 1948 — 620

BOOK REVIEWS by Kenneth W. Haemer, Edgar M. Hoover, Edwin G. Olds, Paul S. Olmstead, George J. Stigler, Alan E. Treloar — 597

VOLUME 43 NUMBER 244

American Statistical Association

Organized November 27, 1839
Incorporated 1841

The American Statistical Association is a scientific and educational organization. Its membership is not confined to professional statisticians but includes economists, business executives, research directors, government officials, university professors, and other persons who are seriously interested in the application of statistical methods to practical problems, in the development of more useful methods, and in the improvement of basic statistical data. Engineers, mathematicians, biologists, actuaries, sociologists, psychologists, and representatives of many other professions are included in the membership of the Association.

Regular membership	$ 8.00*
Student membership	3.00
Introductory membership (for the first dues payment of applicants under 30 years of age)	3.00
Members subscription to Biometrics	2.00
Associate membership in the Biometrics Section	3.00
Contributing membership	25.00

Subscription rate, $8.00 per annum. Prices for back issues available on request.

Additional information about the Association and membership application forms may be secured from the Secretary, 1603 K Street, N. W., Washington 6, D.C.

The Editors welcome the submission of articles and notes for possible publication in the JOURNAL. Manuscripts should be sent to the Editor, Journal of the American Statistical Association, 1603 K Street, N. W., Washington 6, D.C. (1) Send two or more copies of your manuscript (preferably in separate envelopes, to avoid danger of loss). (2) Leave one half of the first page of your manuscript blank, to be used for instructions to the printer. (3) Attach to the front of your manuscript a fifty word outline or summary. Authors who wish suggestions about the preparation of manuscripts and charts or information about editorial policies should address their inquiries to the Editor. Book reviews should be sent to Oscar K. Buros, Rutgers University, New Brunswick, New Jersey.

---

* $4.00 of which is for one year's subscription to the *Journal* of the American Statistical Association, and 75 cents for one year's subscription to *The American Statistician*.

J
S

Nu

T
sho
As
the
the
1
2
3
4
5
6

Int
app

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION

Number 244 DECEMBER 1948 *Volume 43*

COMMERCIAL USES OF SAMPLING*

J. Stevens Stock and Joseph R. Hochstim
Opinion Research Corporation, Princeton, New Jersey

To illustrate how a commercial research agency meets its sampling problems, eight surveys of widely varying types are briefly described.

I

The purpose of this paper is to report on some of the commercial uses of sampling. We shall describe a series of actual examples which show how a commercial research agency meets its sampling problems. As in all sampling problems, the nature of the research objective and the facilities at hand are the all-important conditions which determine the approach.

1. Most commercial agencies must be prepared to tackle an extremely wide variety of sampling problems. No one sampling procedure or principle is applicable in all, or even in a majority of, cases.
2. The actual budget is frequently not always commensurate with the difficulty of the problem. Often it is necessary to do the best job possible under a very restrictive budget, even though a larger budget would allow for much greater dollar efficiency.
3. Most commercial research problems are geared into current-policy determination. Consequently, research efficiency must sometimes be sacrificed for speed.
4. Although there are important exceptions, commercial research problems are for the most part nonrecurrent. Time series and repetitive studies are conducted, but are relatively small in number.
5. Many research problems must be tackled with limited source material because of the specialized nature of the problems.
6. Many commercial research studies must cover unusual, difficult-to-define universes of people, such as the purchasers of a certain commodity, all chemists in the United States, all people who will vote in the next election, etc.

* Presented at the 25th Session of the International Statistical Institute, a constituent part of the International Statistical Conferences, Washington, D. C., September 6–18, 1947. This paper will also appear in the Proceedings of the Conferences, which will be published in the near future.

7. Commercial research agencies often find it necessary to study attitudes and opinions which have highly emotional, personal, or moral implications in the respondent's mind.
8. Very often, even in large national studies, the commercial research agency must make detailed analyses for relatively small areas within the country. This makes it necessary to maintain extremely large part-time staffs widely distributed throughout the nation.

Regardless of any pet theories, the sampler has to consider the following:

1. The problems the user of the data needs to have answered.
2. The actions likely to be based on the results of the survey.
3. The degree of precision required.
4. The size of the budget.

In any case, the sample design chosen should provide the user of the data with maximum information at minimum cost.

II

Our first example involves the sampling of purchasers of alcoholic beverages. It is illustrative of the quick, nonrecurrent survey of a difficult-to-define universe on a subject that for some respondents has moral, personal, and political implications. In one particular state a law had been passed increasing the liquor tax. The client, a distilling company, wanted to know who was getting the blame from purchasers of hard liquor for the increase in prices—the state government, the distillers, or the distributors. The answer to this question would determine the direction the client's public-relations effort would take and the issues on which it would concentrate.

Early testing of the questionnaire indicated that people were unable or unwilling to give accurate answers as to their purchases of alcoholic beverages, except at the time of purchase. The client, however, was interested in directing his public-relations effort at the purchasers of his product so that it would have the greatest impact. It was necessary, therefore, to design the sample so that the number of respondents was approximately proportionate to the amount of total liquor sales throughout the universe.

In the state being surveyed packaged alcoholic beverages are sold only in state liquor stores. Data on liquor sales were available for the past year for different districts of the state, for individual towns in each county, and for individual stores. The sample was designed as follows:

The number of interviews with purchasers in the various districts

was computed in proportion to the amount of liquor sales in each district. The two largest cities of the state were removed from their districts and treated as separate units. Interviews were allocated to these two cities separately according to the amount of their own liquor sales.

Each district was divided into three groups of counties: those with high per-capita consumption, those with medium, and those with low per-capita consumption.

Within each of these three groups of counties the towns were stratified by per-capita volume of liquor sales. One sample town was selected from each substratum. The number of interviews with liquor purchasers allocated to each sample town was in proportion to the total volume of liquor sales of the substratum from which the town was selected. Within each town a list of stores was prepared and arranged in rank order according to volume of liquor sales. Stores were then selected within each sample town systematically with probability of selection proportionate to sales volume.

Interviewers were instructed to stand outside the stores selected and to interview every *n*th person leaving the stores until their allocated number of interviews was complete. In order to assure proper distribution by time of day and day of the week, interviewers were instructed to interview half of their respondents on Friday afternoon and Saturday, and the other half on other days of the week. In addition, the interviewers were instructed to fill half their their quotas before 4:00 P.M. each day and half of them after 4:00 P.M. each day.

III

Our next example is typical of a study of the community relations of an industrial firm.

The largest employer in a Midwestern town of about 100,000 population felt that his company had done a good job of public and industrial relations. But in spite of all its efforts the company had a strike. The executives wanted to know very quickly what could be done to correct an apparent hostility of the citizenry of the community. The problem of the survey was to find what the people in the community were thinking about the company and its employees and what the company could do in the opinion of the community to make the plant a better place to work.

The 1940 Census Block Statistics, which were available for this city, were brought up to date in numbers of dwelling units with the aid of the city planner's data on new construction. All blocks were then

stratified first by census tract, then by average rent. Within each of these strata they were arranged by the number of dwelling units per block. The sample blocks were selected by systematic random sampling with probability of inclusion proportionate to the number of dwelling units in each block.

Within each block the interviewer was given a starting point and instructed to go in clockwise direction and to interview every *n*th dwelling unit in the assigned block until the quota for that block was filled.

In cases of noninterview because of refusal or because no one was at home, interviewers were told to go to the next dwelling unit, but to continue the count from the dwelling unit originally designated. In order to minimize the bias resulting from the omission of not-at-homes and to give more adequate opportunity for getting employed workers in the sample, interviewers were required to do their interviewing in the evening and over weekends.

Quotas covering the whole city by sex and age were given to the supervisor. By cumulative daily tallies of the interviews returned, the supervisor was able to watch the representativeness of the survey. He assigned varying quotas to the interviewers for the selection of individuals within the sample households when certain age-sex groups in the population appeared to be under-represented according to the over-all quota.

Note: Recent experience with this combination of areal and quota techniques indicates that much of the interviewer's tendency toward bias is eliminated by the systematic pre-selection of each dwelling unit in which an interview is to take place. The use of quotas for the selection of an individual within each household allows the interviewer very little free choice.

IV

The next example is a survey to measure the attitudes of social-science teachers toward basic principles of free-enterprise economics. The particularly difficult problem in this case was to get a cross section of a highly specialized population, namely, college and high-school teachers of government, economics, sociology, and history.

The sample of high-school teachers was allotted by states according to data on high-school enrollment published by the Office of Education. Interviews were allocated to cities and towns in proportion to the population distribution of urban places within each state. The selection of the 160 cities and towns within states and size-of-town groups was determined by the location of the company's resident interviewers.

Note: At first glance, it would seem the selection of towns to suit the convenience of the resident interviewers is quite unrigorous. In this case, however, it is an excellent means of selecting a random sample of towns. Since there is no conceivable correlation between the methods by which part-time resident interviewers are recruited and any attitude characteristic of the social-science teachers being studied, we can say that we have selected our sample at random.

It is a common practice in commercial research firms (and others as well, we presume) to base the selection of localities for the cross section on the location of the resident staff. Many suspect this technique of bias. However, as long as the investigator can guarantee that the method of recruiting is in no predictable way connected with the problem studied, we can be assured that the selection of localities is as random as if it had been based on the use of random numbers or some game of chance.

Most commercial research agencies that maintain large part-time resident interviewing staffs recruit them in proportion to the distribution of the population by region, state, and size of town groups within states. In this case the total maximum number of resident interviewers available was approximately 1,000 located in more than 600 places.

As for college teachers, a complete list of all colleges was prepared by type—senior, junior and teacher colleges—and was arranged by states. The number of teachers was recorded for each college, and the colleges were put into two groups: those with a teaching staff of 200 or more and those with fewer than 200 teachers. Every *n*th college was selected for the sample.

Since in neither case was there available a total count on any social or economic characteristic of teachers of social-science subjects, it was impossible to assign any quotas by characteristics. It was necessary therefore, to insure a strict randomization of the respondents. Accordingly, interviewers were instructed to obtain a complete list of social-science teachers for the sample colleges and all high schools in the towns selected. The interviewers were further asked to alphabetize this list and to interview the first n teachers starting with a given letter of the alphabet. Interviewers were given different starting letters.

A separate part of this survey illustrates the use of mail questionnaires in studying a highly specialized universe. A large industrial company had conducted for social-science teachers a course of several weeks in some of the social problems faced by business and the techniques used in solving these problems. The purpose of this study was

to measure the impact on teachers of impressions obtained in this course.

A complete list of all participants in the course was obtained. Post cards were mailed to participants asking them to state by return mail whether they would be willing to answer a mail questionnaire.

A great majority of the teachers expressed their willingness to answer the mail questionnaires. When the returns were counted, however, slightly less than 50% of all participants actually mailed questionnaires.

Since it was felt that the returned sample might not be representative and that those teachers who had not returned the mail questionnaires might be different in their characteristics and attitudes, it was decided to take a separate sample of nonreturners by personal interviews.

The industrial company that gave the course was located in a very large Midwestern city, and about 40% of the teachers attending the course lived within an area of 15 miles from the plant city. It was decided that resident interviewers were to telephone all social-science teachers included in the original list who lived within a 15-mile radius of the plant city. They were to inquire whether the teachers had returned the mail questionnaires and if not, to make an appointment for a personal interview.

By this method, interviews with approximately one fourth of the teachers who did not answer the mail questionnaires were secured. It was then assumed that those who did not answer the questionnaires and who lived in the plant-city area were representative of all the teachers not answering the questionnaire. Accordingly, the answers to personal interviews were weighted to give proportionate representation to those who had, and those who had not, answered the mail questionnaires.

V

Much commercial research is devoted to studying the style or product preferences of highly specialized groups in the population who are actual or potential customers for a certain commodity. The purpose of such research is to get rough estimates of relative customer preference in order to guide manufacturing and sales policy.

A survey of this kind was recently conducted to determine style preferences for a particular piece of furniture. In order to set up as nearly as possible a realistic buying situation, the study was restricted to three separate groups:

1. Well-to-do respondents were questioned on styles in high-priced items.

2. Average and medium-well-to-do respondents were questioned on medium-priced items.
3. Poor respondents were questioned on styles in low priced items.

To simplify the approach and meet the restrictive budget, interviews were conducted only in towns with population of 50,000 or more. All towns having 50,000 or more population were stratified by geographic sections. Interviews were allocated to cities and towns in proportion to their population in groups of 15 or multiples of 15. In the case of small cities, this amounted to arraying them by population size and selecting an ordered random sample of cities. However, instead of selecting every *n*th city, we selected every *n*th person, so that the probability of selecting any one city was determined by the chance of selecting any one person living in the city. Therefore, the probability of selecting each city was proportionate to its population size.

Note: A common practice of commercial research agencies is to assign interviews in groups of 10 or 15, in order to give each interviewer an assignment large enough to make it worth his while to work on the study and small enough so that he can finish his interviewing in the time limits set for the study.

Within each of the cities selected a relatively large primary sample of 100 to 300 blocks was chosen by selecting every *n*th block from the published block statistics. From these large samples of blocks smaller samples of blocks were selected for assignment.

Note: This practice of selecting relatively large primary samples of blocks, then stratifying them by section of the city, rent, and numbers of dwelling units, for further selection of smaller samples of blocks amounts to *double sampling.* It appears to be a very useful practice where small samples of blocks from relatively large cities are needed. The very practical reason for this kind of double sampling is to save on the burdensome clerical time that would be required to stratify all the thousands and tens of thousands of blocks in a large city. It is necessary only to stratify the primary sample of 100 to 300 blocks.

Within each city the primary sample of blocks was divided into three broad groups: high-rent, medium-rent, and low-rent blocks. In the large cities in the sample, where as many as three groups of 15 interviews were assigned, three interviewers were assigned to sample blocks corresponding to each of the three rental groups. In smaller cities, where only one or two sets of 15 interviews were assigned, interviewers were required to interview in sample blocks from (correspondingly) only

one or two of the rental types. Thus, no attempt was made in smaller cities to get a cross section of each city since no individual city analyses were required.

Note: The reason for this assignment procedure was that a different questionnaire was used for each type of respondent—the well-to-do, the average, and the poor. Because of the complicated nature of the interview and the danger of confusing the questionnaires, only one type of questionnaire was assigned to each interviewer.

Blocks within each of the three broad rental groups in each sample city were selected by stratifying the primary sample of blocks for each broad rental group into three strata according to rent. From each of these three rental substrata, three blocks were selected at random with probability of selection in proportion to the number of dwelling units on each block. Thus, each interviewer was assigned one type of questionnaire in one broad rental group. Further, he was assigned three blocks from each of his three rental substrata (9 blocks in all). The interviewer was instructed to take five interviews from each set of the three blocks in each rental substratum. He was also instructed to go around as many of the three blocks as necessary to contact respondents owning the particular piece of furniture or saying they were in the market for it. If the respondent did not meet this qualification, the interview was discontinued and contact was made with the next dwelling unit. Wherever the respondent owned such a piece of furniture or was in the market for one, the interview was completed.

Various styles of furniture of different price range were judged by respondents (higher-priced pieces by respondents living in the higher-rent blocks, lower-priced pieces by respondents in the lower-rent blocks). On a few pieces of medium price, however, we measured preferences in all three groups for purposes of experimental control.

VI

A frequent problem for any research agency is the high expense of travel from interview to interview. The next example illustrates a problem in studying farmers' opinions, one of the worst situations from the standpoint of interviewer travel cost.

During war time, government regulations determined prices and conditions under which much farm produce could be sold by farmers to food manufacturers and processers. There was considerable resentment on the part of the farmers against these regulations. There was also resentment against the manufacturers who in the opinion of some

farmers manipulated these regulations to their own advantage and to the farmer's disadvantage.

The purpose of the survey was to study farmers' sentiments on this issue. It was decided to restrict the study to seven counties in a Midwestern state—an area where the client did most of his purchasing.

Interviews were allotted to each county in proportion to number of farms, but in order to save cost, interviewing was limited to every other township.

State highway maps were obtained on which square-mile sections within each sample township were numbered.

The number of sections within each sample township was recorded, and a cumulative total of these sections was obtained. Then the average number of farms per section was computed. In each county the desired number of sections was obtained by dividing the assigned number of interviews by the average number of farms per section. In order to concentrate interviewing still more, two adjacent sections were combined into one double section. The sampling ratio was then obtained by dividing the total number of sections in the sample of townships by the number of double sections wanted. An ordered sample of every nth double section was selected. Interviewers were instructed to interview every farmer in the double sections selected, making callbacks wherever the farmer was not found at home.

Note: This concentration of interviews in double sections is, of course less efficient in a statistical sense than a sample of single sections or individual farms. In the information-per-dollar sense, however, it is undoubtedly much more efficient. Moreover, it greatly facilitates callbacks.

VII

Frequently firms desire to make opinion studies among their customers, or associations want to poll a cross section of their members. Often in such cases the client firm or association can supply a complete list of the universe to be studied. The sampling problem is then usually very easy; we stratify the names on the list as efficiently as possible and select every nth name.

Shortly before the end of the war, the planning board of a metropolitan area was concerned with the potential post-war employment in the area. During war time the population of the area had increased very considerably, and it was important to make an estimate of the number of people for whom employment might be provided in the years to come. The budget for the survey allowed interviews with 1,500

employers in the area. A complete set of cards showing each firm, its industry group, and the numbers of its employees was made available by the client. In the area to be studied there were approximately 750 firms having 50 employees or more, which accounted for about 54% of all employees in the region. It was decided to include all these 750 firms in the sample.

Of the remaining 15,000 firms having fewer than 50 employees every 20th one was selected.

Thus, approximately equal numbers of firms having more and fewer than 50 employees were selected for the sample. By full coverage of firms that were more important in terms of employment, the statistical error in projecting the average number of employees was reduced considerably.

Note: This example illustrates the common case in which certain segments of the population to be studied contribute such high variability to the final statistical estimate that it is efficient to sample disproportionately from different strata. This amounts to a very simple application of what is commonly called "Neyman sampling."

The cards for employers of 50 persons or less were stratified by detailed industry groups, and were arranged within each group in ascending order from the smallest to the largest firms. Every 20th card of each industry group was then selected, carrying over the count from one industry group to the next.

VIII

The next survey, which uses quota sampling, illustrates a very important and useful technique for studying many different and widespread populations at the same time.

In this instance the purpose of the survey was to bring out pertinent facts bearing upon the public-relations program of a particular industry. A variety of issues was to be covered, and each issue was to be explored thoroughly.

The research problem was to describe in detail the attitudes of many different "publics" toward this industry:

1. Manual workers employed in this particular industry.
2. Residents of each of the major areas (each was analyzed separately) in which the industry is concentrated.
3. Separate cross sections of the general population in seven geographic regions.
4. A national cross section of the continental United States.

Note: More and more in public-relations research it has been found useful to study the "concentric circles" of populations toward which the public-relations effort is being directed. This puts great premium on developing efficient sampling techniques for overlapping populations. In such cases ingenious and workable methods are needed to take fullest advantage of the additional efficiency gained through making each successive sample contribute to the sample of the next concentric population.

The national cross section was made up of interviews allotted to approximately 450 localities in proportion to the population of geographic regions, states, and of size-of-town groups. Each interviewer was assigned quotas by sex, age, race, employment status and occupation. These quotas were based on 1940 Census data adjusted wherever possible to more recent Government estimates.

In four of the seven geographic regions, the national cross section sample was adequate. In the other three, additional interviews were taken to supplement those obtained in the national cross section. These extra interviews were assigned to the sample localities in proportion to the distribution of the general public in these regions.

Major industrial areas were defined as all counties in which a minimum of 3% of all employed men were engaged in the clients industry. These counties were stratified according to population and according to per cent of men employed in the industry. These substrata were set up so that the aggregate populations of the counties in each were approximately equal. A sample of counties was then drawn at random from each stratum. General-public interviews in the national or regional samples which were taken in counties where more than 3% of the employed men were engaged in this industry were included in the sample of major industrial areas. Consequently, the number of interviews allocated to counties in each stratum was reduced by the number of interviews already taken in this stratum in other samples. Since there was considerable overlap between the different areas analyzed, very substantial savings were effected by taking advantage of the duplications. Quotas were assigned by sex, age, race, employment status, and occupation in proportion to the distribution of such characteristics in each major industrial area.

Interviews with industry employees were taken in the sample counties in the major industrial areas in proportion to the employment in the industry in those areas. The remainder of the industry employee sample was allocated among the national sample localities in proportion to the employment in the industry in those localities.

Note: Quota sampling has been widely condemned in recent years. Principally, its validity can be questioned on the underlying inverse logic on which it is based—namely:

1. A representative cross section will reflect known characteristics accurately;
2. Therefore, a sample which reflects known characteristics accurately is a representative cross section.

In spite of this dubious logic, it can be shown that under certain not uncommon conditions quota sampling may have the highest cost efficiency of any method now known.

IX

A very interesting and ingenious application of an older technique, which has been largely abandoned in the past ten years, has recently been reported by Benson, Young, and Syze.[1]

This technique, now known as pinpoint sampling, was used by the American Institute of Public Opinion in connection with the 1944 presidential election.

Past election results for five particular states were found to represent a remarkable reflection of national election outcomes for all two-party contests since 1916.

Within each one of these five states, eight precincts were selected for the sample. The precincts within each state were chosen in such a way that, "giving equal weight to each, the average of the percentages of the vote of either major party was virtually the same as the election result for the state." This was found to have been true in three previous statewide elections.

Within the precincts selected interviewers were instructed to interview in every—or every other—dwelling unit, depending on the number of voters in the precincts. The interviewers still retained the freedom of selecting respondents within the family group, except for the limitation of obtaining over-all quotas for sex and age.

Interviews were completed in about one half of the households in the sample election districts. Slightly more than one fourth of the potential voters were interviewed.

The results show a deviation of 1.9% from the actual national vote. Part of this deviation was due to the selection of states and part to the selection of precincts. Benson, Young, and Syze concluded from the

---

[1] Edward G. Benson, Cyrus C. Young, and Clyde A. Syze, "Polling Lessons from the 1944 Election," *The Public Opinion Quarterly*, Winter 1945–46. The authors are indebted to Messrs. Benson, Young, and Syze for their permission to present their findings.

results that improvements could most profitably be made by enlarging the number of sample areas and taking a smaller number of interviews within each area.

We have here a case of sampling by purposive selection.

Note: The use of the method of purposive selection has been roundly condemned by many writers on the subject of sampling. Gini and Galvani in an article in *Annali Di Statistica* in 1929 showed empirically that this method was very inefficient. Later Neyman in his now famous article, "On the Two Different Aspects of the Representative Method," *The Journal of the Royal Statistical Society*, 1934, confirmed theoretically Gini's and Galvani's findings.

Neyman, however, did express the opinion that under certain very rigid conditions purposive selection might be an efficient method of sampling. One of those very rigid conditions was that the technique might "work" when the controls were very highly and linearly correlated with the statistics to be measured. Although Benson, Young, and Syze do not discuss this point, it appears that this condition has been met. Therefore, it seems reasonable that the method of pinpoint sampling may have very important possibilities in predicting election results.

X

Although the problems already presented indicate the great variety of situations a commercial research agency has to meet, there are many more entirely different sampling and interviewing situations that cannot be covered in detail here.

Some of the sampling techniques described vary according to the particular survey situation. In a survey taken in a town and in its surrounding rural area, for neither of which block statistics were available, the sample was designed by drawing routes on a town map and state highway maps. The routes were drawn in such a way as to include different sections in the town as well as the different types of roads in the survey area. Interviewers were to interview in every *n*th house on the right side of the route designated until their quota was completed. Different starting points were chosen for the routes, some leading from the town towards outlying areas, others starting as far as 15 miles from the town and leading toward it.

Another example of a special situation: Even though interviews are generally done personally with only *one* respondent in any one household, there are cases where it is best to interview husband and wife together in order to approximate a family buying situation (for example, when a survey is made on intended purchases of refrigerators).

Even though most questioning is done by personal interview, there

are situations where telephone interviews are satisfactory (for example, a survey with a short and strictly factual questionnaire and a specialized universe all members of which are telephone subscribers).

In other cases, such as in "in-plant studies," self-administered group interviews are used with a sample including every nth employee from the payroll.

In situations where respondents might feel that their statements could place them in some form of jeopardy, it may be necessary to interview on the street to allay any possible suspicion that respondents might be traced.

In cases of a management sample, such as sales managers or public-relations directors of companies with assets of more than $5,000,000, it may be necessary to sample from such sources as *Moody's* and *Standard and Poor*, and from individual state directories, supplemented by whatever information is available through local chambers of commerce.

In a survey with stockholders, where the composition of the universe was unknown, information was obtained on the characteristics of the types of people who are most often owners of stocks by including a question on stock ownership in other large national surveys. The survey with stockholders could then be restricted to these types, with the realization, of course, that the unusual types of stockholders would not be included.

There are numerous other situations—radio surveys, magazine surveys, advertising studies, and many others—which require various approaches suited to their special needs.

Looking back at the examples given, we realize that there are many other possible sampling approaches to these problems. The purpose of this paper is to report, not to analyze. These were the approaches that were actually used. They are presented here to illustrate a few of the many commercial uses of sampling.

VARIATION OF THE FREQUENCY OF FATAL QUARRELS WITH MAGNITUDE

LEWIS F. RICHARDSON
Hillside House, Kilmun, Argyll, Britain

A record of wars during the interval A.D. 1820 to 1945 has been collected from the whole world, and has been classified according to the number of war-dead. The smaller incidents have been the more frequent, according to a fairly regular graph which can be extended to quarrels that caused a single death.

INTRODUCTION

ANYONE WHO TRIES to make a list of "all the wars" (e.g. Wright [18] p. 636) encounters the difficulty that there are so many small incidents, that some rule has to be made to exclude them.

From the psychological point of view a war, a riot, and a murder, though differing in many important aspects, social, legal, and ethical, have at least this in common that they are all manifestations of the instinct of aggressiveness. This Freudian thesis has been developed by Glover [5], by Durbin & Bowlby [3], and by Harding [6]. There is thus a psychological justification for looking to see whether there is any statistical connection between war, riot, and murder. I dealt with this problem in 1941 by forming the inclusive concept of a "fatal quarrel," and then classifying the fatal quarrels according to the number of quarrel-dead [12]. By a *fatal quarrel* is meant any quarrel which caused death to humans. The term thus includes murders, banditries, mutinies, insurrections, and wars small and large; but it excludes accidents, and calamities such as earthquakes and tornadoes. Deaths by famine and disease are included if they were immediate results of the quarrel, but not otherwise. In puzzling cases the legal criterion of "malice aforethought" was taken as a guide.

The record of the number killed in a particular war is often uncertain by a factor of two. The meaningful part of the record can be separated from its uncertainty by taking the logarithm to the base ten and rounding the logarithm off to a whole number, or to the first decimal,

according to the quality of the information. For simplicity I have lumped together the deaths on the opposing sides of the quarrel. The *magnitude* of a fatal quarrel is defined to be the logarithm to the base ten of the number of people who died because of that quarrel. The magnitude will be denoted by μ. The range of magnitude extends from 0 for a murder involving only one death, to 7.4 for the Second World War. Other well-known wars had magnitudes as follows: 1899–1902 British versus Boers 4.4; 1939–40 Russians versus Finns 4.83; 1861–65 North American Civil War 5.8. The magnitude of a war is usually known to within ± 0.2; so that a classification by unit ranges of magnitude is meaningful. These ranges have been marked off at 7.5, 6.5, 5.5, 4.5, 3.5, 2.5, 1.5 and perhaps at 0.5 and -0.5. Every fatal quarrel must lie inside one of these cells; it cannot lie on a boundary, because the antilogs of the boundary μ are not integers.

This abstract framework becomes of interest when the facts are sorted into it. As far as I have been able to ascertain, the historians have never sorted their facts by the scale of magnitude. The chief obstacle to any scientific study of wars-and-how-to-avoid-them, is each nation's habit of blaming other nations. National prejudice can however be avoided by taking the whole world as the field of study; and by taking a time-interval longer than personal memory. For this reason I have made a search for the records of fatal quarrels in the whole world since the beginning of A.D. 1820. The task has been long; and the results are here presented in brief summary. For magnitudes greater than 2.5, the facts were mostly obtained from works on history. For magnitudes less than 0.5 they were taken from criminal statistics. For magnitudes between 0.5 and 2.5 the information is scrappy and unorganized; what there is of it suggests that such small fatal quarrels were too numerous and too insignificant to be systematically recorded as history, and yet too large and too political to be recorded as crime.

It is important to know whether the facts have all been gathered. The best evidence is provided by the progress of search. I began in the year 1940. At first my collection grew rapidly, then slowly; then various revisions caused its totals to oscillate slightly. The collection appeared to be sufficiently complete to warrant the publication of a summary, which appeared in *Nature* of 15th November, 1941. Afterwards the publication of Quincy Wright's list of wars ([18] Appendix XX) provided a stimulus to further enquiry, which involved the consultation of some seventy history-books. In the following table three stages of my records are compared. All refer to fatal quarrels which ended from A.D. 1820 to 1929 inclusive.

Date and place of record	Ends of range of magnitude			
	$7\pm\frac{1}{2}$	$6\pm\frac{1}{2}$	$5\pm\frac{1}{2}$	$4\pm\frac{1}{2}$
	Numbers of fatal quarrels recorded			
1941. *Nature* of Nov. 15	1	3	16	62
1944. J.R.S.S. 107, 248	—	—	—	63
1947. Here.	1	3	20	60

It is seen that there has been a sort of convergence, which makes the present numbers worthy to be discussed. That is for $\mu > 3.5$. For smaller incidents in the range $2.5 < \mu < 3.5$ my collection has continued to grow; the sign $\geqq$ is therefore prefixed to its totals.

The world-total for the murders was obtained from the murder-rates of 17 countries. They were weighted by populations. The weighted mean rate for the world was found to be 32 murders per million per year. This is not as definite as the statistics of wars, because the criminal statistics of China, Africa, and South America (except Chile) were not available. But even if that estimate were as much as three times wrong, which is incredible, it still would give significant results in Figures 0 and 1, because they relate to such enormous ranges. The evidence which I had concerning time-changes in the murder-rate was scrappy and conflicting; so that provisionally I took the rate to be constant. The mean population of the world over the 126 years from A.D. 1820 to 1945 was computed from a publication by Carr-Saunders [1A] and was found to be 1.49×10^9. So the world-total of murders was computed as $1.49 \times 10^9 \times 32 \times 10^{-6} \times 126 = 6.0 \times 10^6$.

A revision based on the murder-rates of 21 countries, gave a mean rate of 37 instead of 32, and a world-total of 7 instead of 6 million. The following diagrams and tables were based on the earlier estimate and have not been altered.

A CONSPECTUS WHICH SUITS THE WARS

The number of fatal quarrels has been counted, or estimated, in unit ranges of magnitude, with the following results for the world as a whole. The date of a quarrel is taken to be that of the termination of hostilities. The interval is A.D. 1820 to 1945 inclusive.

Ends of range of magnitude	$7\pm\frac{1}{2}$	$6\pm\frac{1}{2}$	$5\pm\frac{1}{2}$	$4\pm\frac{1}{2}$	$3\pm\frac{1}{2}\cdots$ Murders
Observed number of fatal quarrels	2	5	24	63	$\geqq 188 \cdots 6 \times 10^6$

It is desirable to show all these facts together on a single diagram. If the number of fatal quarrels per unit range of magnitude were taken

as the ordinate, then either the wars would show, or the murders, but not both. When however the logarithm of the number of fatal quarrels per unit range of magnitude is taken as the ordinate, then all parts show equally well. Each murder presumably involved a small number of deaths, such as 1, 2, 3, or 4, so that the corresponding logarithms were 0.000, 0.301, 0.477, or 0.602. How to group them was not obvious. As a first expedient I extended to the murders the system of classification which had been found to suit the wars, namely by unit ranges of magnitude cut at 7.5, 6.5, 5.5, 4.5, 3.5, and so on by extension to 0.5 and -0.5. Murders involving 4 or more deaths were regarded as negligibly rare. The ordinate f is defined to be

$$f = \log_{10} \left\{ \frac{\text{number of fatal quarrels}}{\text{corresponding range of magnitude}} \right\}.$$

This rough and ready scheme gives the conspectus shown in Figure 0. It is the type of diagram with which the author began, and it has now been asked for by a referee. A dotted curve has been drawn across the gap to suggest that there is statistical continuity between the wars and the murders.

THE FAULTS OF FIGURE 0 AND HOW TO AMEND THEM

In order to match the wars, the murders have been enclosed in a unit range of magnitude, as shown at A in Figure 0. But the half of this range, from $\mu = -0.5$ to μ just less than zero, is necessarily empty. Would it be more reasonable to regard the murders as enclosed in a half-unit range? If so the line representing them in Figure 0 would be halved in length. Also the number of murders per unit of magnitude would be doubled. So the representative line would be raised by $\log_{10} 2$ into position marked B. Worse ambiguities occur if we wish to show separately the number of murders each causing one death. Let there be M of them. To what range of magnitude ought they to be attributed? The range must contain the point $\mu = 0$, and must not contain the point $\mu = \log_{10} 2$, but otherwise is unspecified. Yet M has to be divided by this indefinite range in order that the reckoning may be "per unit range of magnitude". Figure 0, though suitable for the wars, is too vague for quarrels that caused only one death.

It may be thought that the root of the trouble is the reckoning "per unit range of" anything, and that if ranges were abandoned then murders and wars could be compared by plotting the logarithm of their numbers against their magnitudes. Such a diagram would consist of

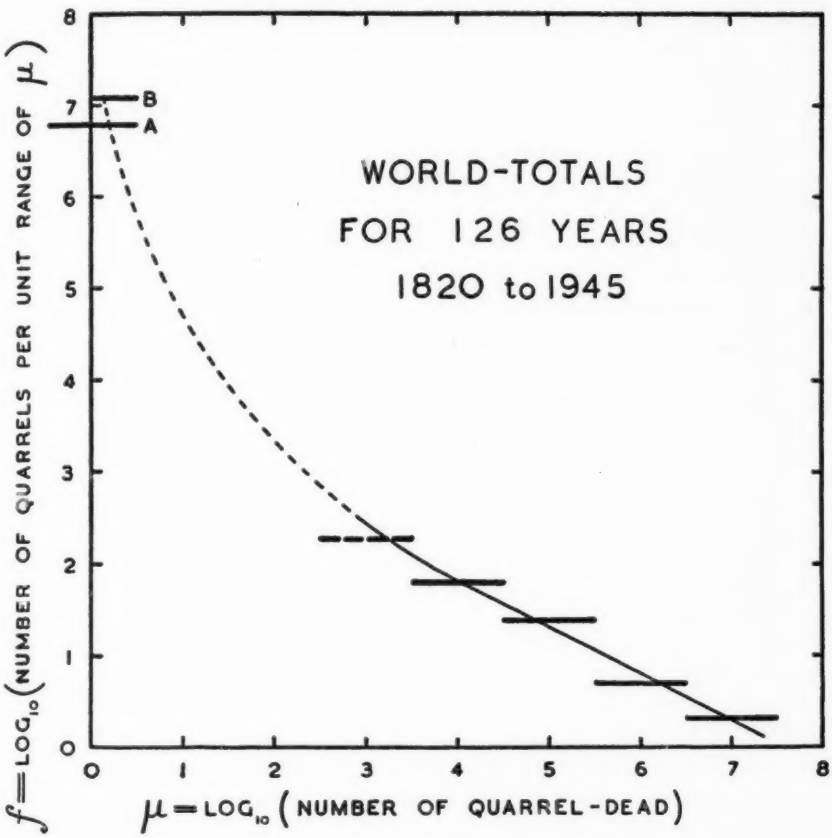


FIGURE 0

points, not segments. There are two objections to this proposal: (i) the magnitudes of wars are imperfectly known and some of them are rounded off to whole numbers; so that there would be many artificially coincident points, (ii) if the magnitudes were better known they would presumably be scattered, so that a point-diagram of wars would present great irregularities, of no interest. For when n exceeds 1000 there is never likely to be any statistically significant distinction between wars involving n and $n+1$ deaths. In brief we must have points for the murders and averages over ranges for the wars, and yet these diverse representations must somehow be compared. The solving idea, which melted all the obstacles, was that the appropriate measure of any range of quarrel-dead is simply the number of integers in that range, and not

the corresponding range of magnitude, nor any other artificial conception.

A DIAGRAM WHICH SUITS THE MURDERS

The logically simple idea is that there were, say $q(n)$ quarrels each involving n deaths. (1)

For example in England and Wales during 1935 and 1936 there were known to the police 186 cases of murder of 213 persons aged over one year. In 74 cases (involving 91 persons) the murderer or suspect committed suicide. In 105 cases (115 victims) 111 persons were arrested. In the remaining 7 cases, involving 7 victims, no arrest was made. Of the 111 persons arrested 16 were executed (Whitaker's Almanacks 1938 and 1939, p. 652). These statements can be interpreted, with only slight uncertainty, by the following distribution.

			totals
$n=1$	2	3	
$q(n)=92$	71	23	186
$nq(n)=92$	142	69	303

The number of deaths, n, connected with the quarrel is here taken to include, along with those of the victims, also those of the murderer or suspect, whether by suicide or execution. It seems permissible, for illustration until better information can be found, to regard the 6 million murders in the world from 1820 to 1945 as the number of fatal quarrels for $n=1$, 2 and 3, and to subdivide this total among its three parts in the same ratios as the total for England and Wales is subdivided. Thus one finds for the world, the following enlarged values of $q(n)$. To take one country as a sample of the world is very unsatisfactory. But it should be noted that the uncertainties of these $q(n)$ are insignificant in comparison with the vast range of Figure 1 where $q(n)$ varies in a ratio exceeding 10^{13}.

			totals	
$n=1$	2	3		
$q(n)=3.0$	2.3	0.7	6.0	millions
$nq(n)=3.0$	4.6	2.1	9.7	

These give three points in the top left corner of Fig. 1, with the coordinates

$\log_{10} n=0$	0.301	0.477
$\log_{10} q(n)=6.48$	6.36	5.85

A GENERAL AND WELL DEFINED CONSPECTUS

This will now be obtained by extending to the wars the method introduced above for the murders.

In the region of wars, say for $n>1000$, the function $q(n)$ is usually zero and occasionally unity; so that a diagram of it would be bristly and unreadable. Therefore for the wars we should take an average of $q(n)$, say $\bar{q}(n)$. As a preliminary to the definition of $\bar{q}(n)$, it must be understood that the fatal quarrels are arranged in order of n. The definition of $\bar{q}(n)$ is then

$$\bar{q}(n) = \frac{\text{number of fatal quarrels in a range of } n}{\text{number of integers in that range of } n}. \qquad (2)$$

Briefly, and somewhat inaccurately, $\bar{q}(n)$ may be called the "average frequency of quarrels per unit range of quarrel-dead." As in all frequency-diagrams so here the choice of the group-range has to be a compromise: if the range is too small, the diagram is spiky; if the range is too large, essential features are flattened out. The question of the best range must be decided by trial.

The diagram of q and $\bar{q}$ as functions of n is the logically simple conspectus of the frequency of wars and murders. This diagram cannot be drawn, because no sheet of paper is large enough to show the facts at both ends of the range. The difficulty is overcome by plotting the logarithms of q, $\bar{q}$ and n. The magnitude μ has already been defined by

$$\mu = \log_{10} n. \qquad (3)$$

Another oft-recurring symbol, ϕ, must now be defined.

Let $\log_{10} \bar{q}(n) = \phi(\mu)$, say, for the wars: while for the murders

$$\phi(\mu) = \log_{10} q(n) \text{ without averaging.} \qquad (4)$$

The last column of the following table shows $\phi(\mu)$, apart from some corrections for grouping, which will be explained on p. 542. The first

Range of magnitude μ	Number of integers in range of war-dead	Observed number of fatal quarrels Years 1820–1945	$\log_{10}$ (number of quarrels per unit range of war-dead)
2.5 to 3.5	2,846	≧188	≧ −1.180
3.5 to 4.5	28,460	63	−2.655
4.5 to 5.5	284,605	24	−4.074
5.5 to 6.5	2,846,050	5	−5.755
6.5 to 7.5	28,460,499	2	−7.153

and last columns of this table provide the abscissa and ordinate of the accompanying diagram (Figure 1), as far as the wars are concerned. It is evident that the wars can be satisfactorily fitted by a straight sloping line. On account of grouping, the wars are represented, not by points

as the murders are, but by horizontal segments. It will be shown later (p. 542) that the $\phi(\mu)$ graph should pass slightly above the midpoints of these segments. That is to say the sloping straight line was actually obtained by a second approximation.

The relation between the ordinates of Figures 0 and 1

The abscissa μ is the same in both diagrams and the ordinates are respectively f and ϕ, defined thus

$$f = \log_{10}\left\{\frac{q(n_1) + q(n_1+1) + q(n_1+2) + \cdots + q(n_1+s)}{\log_{10} n_b - \log_{10} n_a}\right\}$$

where $n_1-1<n_a<n_1$, and $n_1+s<n_b<n_1+s+1$. These extra bits at the ends of the range have to be put in to prevent f becoming infinite at $s=0$.

$$\phi = \log_{10}\left\{\frac{q(n_1) + q(n_1+1) + q(n_1+2) + \cdots + q(n_1+s)}{s+1}\right\}$$

which is definite at $s=0$. By subtraction

$$f - \phi = \log_{10}\left\{\frac{s+1}{\log_{10} n_b - \log_{10} n_a}\right\}.$$

This expression involves an ambiguity, because n_b is loose in a range of unity, and so is n_a. The practical question however is whether the ambiguity is greater than the uncertainty of the observations. It is necessary to distinguish different cases.

Case (i). The range here used in treating the wars is $\mu_0-\frac{1}{2}\leqq\mu\leqq\mu_0+\frac{1}{2}$ where $\mu=\log_{10} n$ and where $\mu_0\geqq 3$. In these circumstances arithmetic shows that

$$f - \phi = \mu_0 + 0.454, \tag{5}$$

reliable to the third decimal. Seeing that the μ of a particular war is often uncertain by ± 0.1, equation (5) is admirably accurate for the purpose.

Case (ii). We may conceivably wish to consider a shorter range $\mu_0-\epsilon\leqq\mu\leqq\mu_0+\epsilon$ where μ_0 is still in the region of the wars. If ϵ is small, but not too small, there is a quasi-limit such that

$$f - \phi = \mu_0 + 0.362. \tag{5A}$$

This was obtained by approximating to the difference-ratio by a derivative, and was confirmed by arithmetic.

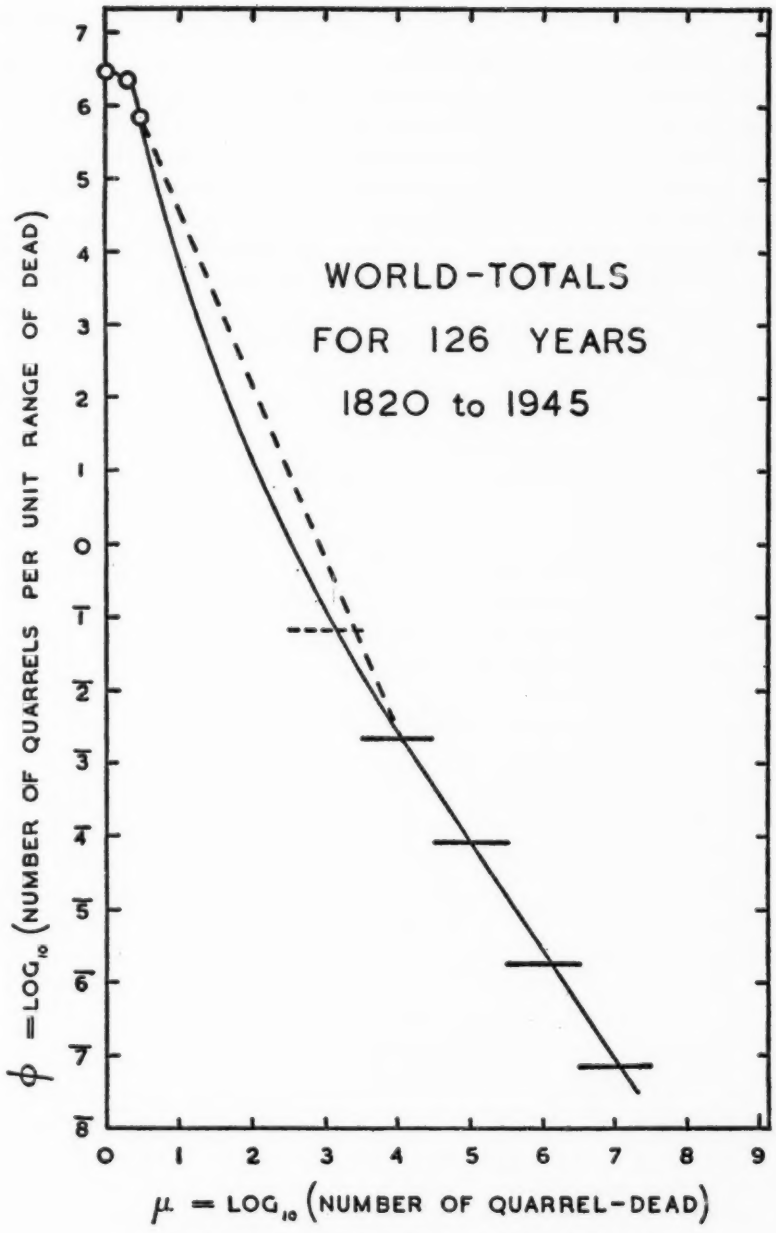


FIGURE 1. THE WHOLE RANGE OF FATAL QUARRELS

THE HORIZONTAL SCALE EXTENDS FROM A SINGLE PERSON ON THE LEFT TO THE POPULATION OF THE WORLD ON THE RIGHT. THE THREE DOTS AT THE LEFT HAND TOP REPRESENT THE MURDERS. THE SHORT HORIZONTAL LINES REPRESENT THE WARS GROUPED BY MAGNITUDES. ONE OF THESE LINES IS DOTTED, BECAUSE IT SHOULD BE RAISED TO AN UNKNOWN HEIGHT.

Case (iii). If there are only a few integers n in the range then the ambiguity becomes troublesome. This is the situation as to the murders. The ambiguity of $f-\phi$ is more than that of some criminal statistics. The fault lies in f, not in ϕ. When $n_1=1$ and $s=0$ it is hardly possible to locate n_a and n_b by any reasonable convention, as already explained on page 526. To revert to Cases (i) and (ii), either of them give the same relation between the slopes of the f- and ϕ-diagrams. As these slopes are in fact negative, it is convenient to write the relation thus

$$-d\phi/d\mu_0 = -df/d\mu_0 + 1 \tag{5B}$$

which shows that the ϕ-diagram is the steeper.

The slope in the region of the wars

This was found, by the principle of Maximum Likelihood, [4 & 10] to be

$$d\phi/d\mu_0 = -1.50. \tag{5C}$$

Interpolation across the gap between the wars and the murders

The dotted horizontal segment for the small wars represents only a lower bound, and otherwise should be ignored. Between the definite end-points of the gap, the mean slope is

$$d\phi/d\mu = -2.38. \tag{6}$$

This is shown by a straight line of dashes.

As to the choice of a curve, there are the general clues of simplicity and of continuous slope. For simplicity a circular arc was chosen; and, for continuous slope, the arc was made tangent to the straight line which fits the wars. It was then found, as an independent confirmation that the slope of the circular arc joins on continuously to that for the murders. Some collateral evidence is provided by banditry.

A SEARCH FOR THE FACTS ABOUT QUARRELS THAT CAUSED FROM 4 TO 315 DEATHS ($0.5<\mu<2.5$)

That there were many small fatal quarrels is shown by a statement made by Quincy Wright [18] in the introduction to his long list of "Wars of Modern Civilization." He remarked that "A list of all revolutions, insurrections, interventions, punitive expeditions, pacifications, and explorations involving the use of armed force would probably be more than ten times as long as the present list, . . . " Wright however did not classify fatal quarrels according to the number of quarrel-dead. Similarly whilst reading histories and police-reports and listening to

radio-news I have noticed allusions to very many incidents less important than wars, but more important than murders. References [2], [7], [11], [15], contain some remarkable examples. However I see no hope of obtaining world-totals of the numbers of such incidents during the 126 years to which Figure 1 relates, and classified moreover by magnitudes in the range $0.5 < \mu < 2.5$. Any factual test of the interpolated portion of Figure 1 will have to depend on smaller samples. The record of banditry in Manchoukuo is comparable on certain assumptions, and that of gangsters in Chicago on further assumptions. These ordered collections of facts, being rare specimens, are correspondingly valuable.

BANDITRY IN MANCHOUKUO DURING THE YEAR 1935

This is of especial interest, because the facts are given in a form which throws light on aggregation for aggression. The following is quoted from the *Japan and Manchoukuo Year Book*, 1938, pp. 692–95.

"At the time of the founding of Manchoukuo in March of 1932 the total number of bandits exceeded 100,000. By September of the same year the number had increased to 210,000 due principally to the subversive activities of Chang Hsueh-Liang's remnant troops who were thrown out of employment following the downfall of the young marshal. Since then, however the number of such bandits has been on the decrease as a result of their suppression by Manchoukuo and Japanese forces.

"Compared with the condition obtaining in 1932 two factors loom in prominence with regard to the bandit situation. Firstly may be noted the actual reduction of bandits as a whole, and, secondly, the shrinkage in size of bandit groups. In 1932 some bandit groups had an actual fighting force of 30,000 men, but at present the average is below 50 bandits per group. The chief cause for the existence of bandits in Manchoukuo is believed to be an economic one, resulting from unemployment."

STATISTICS OF BANDITRY

Size of bandit groups						
1 to 30	31 to 50	51 to 100	101 to 200	201 to 300	301 to 500	>501
Corresponding number of raids						
28,145	4,784	3,864	1,530	455	240	130

The first step towards making a comparison with larger fatal quarrels is to compensate for the unequal ranges of size in the above quotation, by dividing each number of raids by the corresponding range of size. For example $4784/(50-31+1)=239.2$. Any such ratio may be

called the "number of raids per unit range of membership." For the first six groups it runs as follows

938.17 239.2 77.28 15.30 4.55 1.20

and shows plainly that *the smaller incidents were much the more frequent.*

Was a raid like a small battle or like a small war? In other words did the same group of bandits perpetrate many raids? This can be answered by attending to totals. The total number of appearances of bandits can be underestimated from the above table by multiplying the least size in each column by the corresponding number of raids, and comes to 758,371 or more. A moderate estimate is 1.3 million.

But the total number of bandits is stated to have been much less than 210,000. It is evident therefore that on the average the same bandit appeared in several raids. That is to say *a raid resembled a battle rather than a war.*

In the same year and same region the numbers killed in connection with banditry are stated thus: bandits 13,338, suppression troops 1361, civilians 2512. The total dead was therefore 17,211. This is 1.3 per cent of the total number of appearances of bandits.

The facts about banditry are plotted in Figure 2. This is not a $\phi(\mu)$ diagram, yet it is rather like one. For the abscissa, ν say, is $\log_{10}$ of the number of bandits in the group, whereas μ is $\log_{10}$ of the number of quarrel-dead. Again the ordinate, ψ say, is $\log_{10}$ of the average number of raids per unit range of membership, whereas ϕ is $\log_{10}$ of the average number of fatal quarrels per unit range of quarrel-dead. The first group in the data has a membership range of from 1 to 30, which is in a ratio too great for present purposes. Although the first group is shown on Figure 2, it will be ignored in the discussion. The blunted top of Figure 1 has however some resemblance to the horizontal in Figure 2. The other groups are well fitted by a straight line having the slope

$$d\psi(\nu)/d\nu = -2.29. \tag{7}$$

This has a remarkable resemblance to the mean slope, -2.38, across the gap in the world-diagram, as stated in (6). Can this agreement of the slopes be a mere coincidence? Or is it a clue to a general law concerning aggregation for aggression? It is at any rate easily explained by the following two assumptions:—

Firstly let us suppose that the quarrel-dead were on the average proportional to the number of bandits in the group. That is to say that groups continued to make raids until a constant fraction of their members had been killed. The supposition is that

$$n = A \times \text{(number of bandits in group)}$$

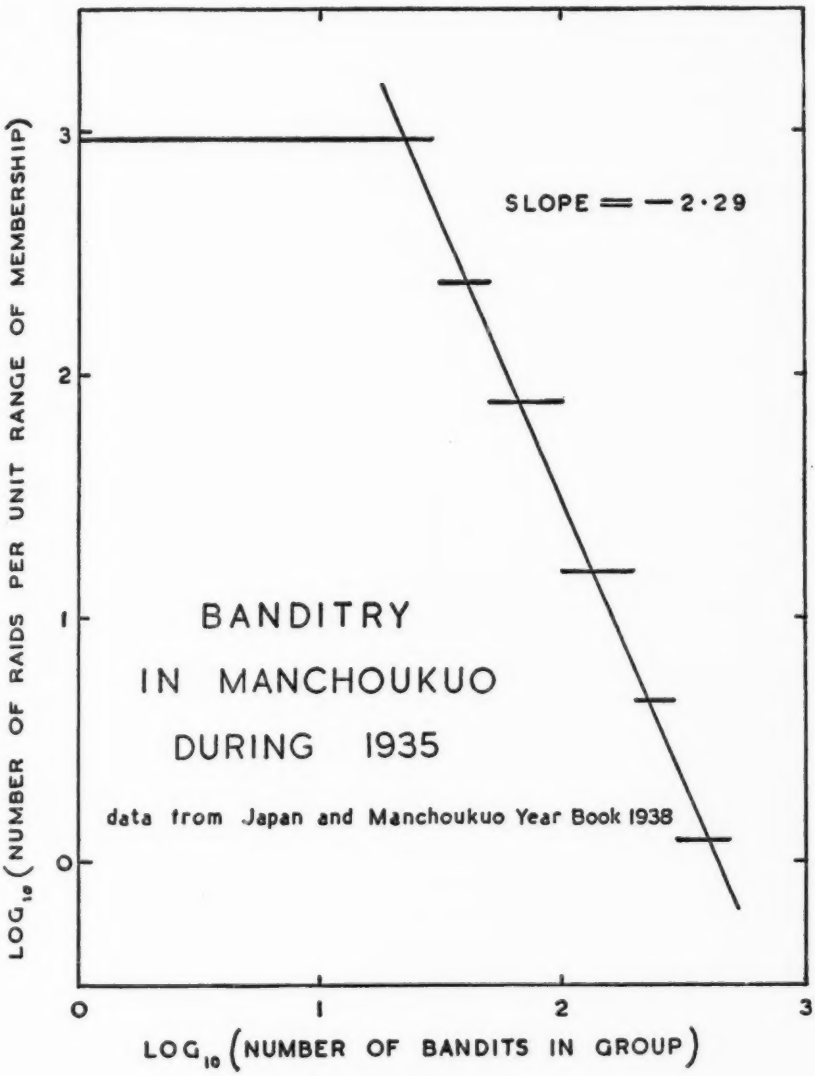


FIGURE 2

where A is independent of the size of the group. Accordingly

$$\mu = \nu + \log_{10} A. \tag{8}$$

It may not be necessary thus to suppose that the percentage of killed was exactly the same for every group, but only to suppose that it had no correlation with the size of the group.

Secondly let us suppose that the number of raids which a group of bandits made, as part of their quarrel with the rest of the community, was not correlated with the size of the group. More precisely the supposition is that:

$$\text{(number of raids made by a group)} = Bq(n)$$

where B is independent of n, so that

$$\psi(\nu) = \phi(\mu) + \log_{10} B. \tag{9}$$

According to these assumptions the $\psi(\nu)$ graph, when slid parallel to itself, horizontally through $\log A$ and vertically through $\log B$, becomes a *local* $\phi(\mu)$ graph. The remarkable agreement of the slopes can therefore be interpreted as meaning that one of the characteristics of aggregation for aggression was the same in Manchoukuo as it was in the world-total.

That broad agreement is interesting; but there remain some doubts or discrepancies in detail. The Manchurian data do not fix A and B, though A can be roughly estimated. For it has been shown that on the average 1.3 per cent of the bandits who appeared in a raid were killed, and that a group on the average perpetrated several raids. The total quarrel-dead, n, may have been about 10 per cent of the number of bandits in the group; that is to say $A=0.1$, so that $\mu=\nu-1$. The mean value of μ on the straight part of the bandit-diagram is accordingly about one. If so, there is a discrepancy with the circular arc on the world-diagram. For at $\mu=1$ the slope of the arc is decidedly steeper, namely $d\phi/d\mu=-3.1$. The Manchurian data therefore suggest that the gap in the world-diagram (Figure 1) should be closed, not by the circular arc, but by the straight segment of slope -2.38 leaving discontinuities of slope at its two ends. I regard that as a suggestion to be remembered, but not to be acted on without further evidence. The shift $\mu=\nu-1$ cuts off most of the horizontal part of Figure 2 and thus increases its resemblance to the top of Figure 1.

GANGING IN CHICAGO

F. M. Thrasher [16] made a study of 1313 gangs in Chicago with a view to their re-direction, or to some other social improvement.

He defined a gang in these words (p. 57): "The gang is an interstitial group originally formed spontaneously, and then integrated through conflict. It is characterized by the following types of behaviour: meeting face to face, milling, movement through space as a unit, conflict, and planning. The result of this collective behavior is the development

of tradition, unreflective internal structure, *esprit de corps*, solidarity, morale, group awareness, and attachment to a local territory."

Thrasher gives (p. 319) a table showing the approximate numbers of members in 895 gangs. The first two columns of the following table are copied from Thrasher, the third column is a deduction, designed to smooth out the disparities of the given ranges. The number of members

Number of members in gang inclusively	Number of such gangs	Gangs per unit range of membership
3 to 5	37	12.3
6 to 10	198	39.6
11 to 15	191	38.2
16 to 20	149	29.8
21 to 25	79	15.8
26 to 30	46	9.2
31 to 40	55	5.5
41 to 50	51	5.1
51 to 75	26	1.04
76 to 100	25	1.00
101 to 200	25	0.25
201 to 500	11	0.37
501 to 2,000	2	0.0013

in a gang is the same sort of quantity as the number of bandits in a group; so the same symbol ν is here used for $\log_{10}$ of the first column. Log_{10} of the third column is here denoted by $\chi(\nu)$ to distinguish it from $\psi(\nu)$, because the number of gangs is not similar to the number of raids, though related to it. Figure 3 shows a graph of $\chi(\nu)$. Apart from gangs of 15 or less, which show again the blunted top, the $\chi(\nu)$ graph is well fitted by the straight line shown, which has a slope

$$d\chi(\nu)/d\nu = -2.30. \tag{10}$$

It is remarkable that this slope is almost the same as $d\psi(\nu)/d\nu = -2.29$ for banditry in Manchoukuo. This agreement strengthens the suspicion that some fairly general tendency concerning aggregation for aggression is revealed by these otherwise scattered phenomena. Hypothetical explanation is again easy. The distinction between the number of gangs and the number of raids would not affect the slope, if it were true that the number, C say, of raids made by a gang was independent of the size of the gang. For then

$$\psi(\nu) = \chi(\nu) + \log_{10} C \tag{11}$$

and the constant $\log_{10} C$ would disappear on taking $d/d\nu$. Yet several features remain unclear. The Chicago gangs were certainly aggressive.

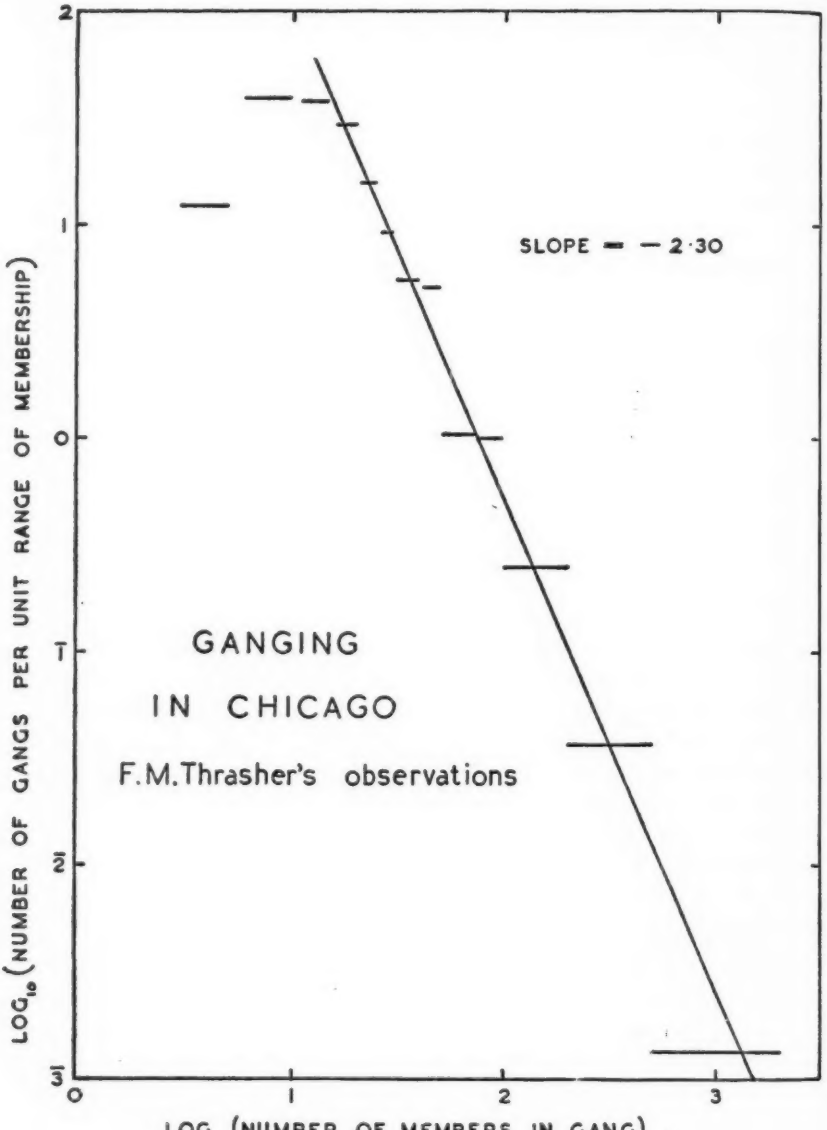


FIGURE 3

Thrasher has a chapter headed "Gang Warfare" and beginning with these words

"The gang is a conflict group. It develops through strife and thrives on war-

> fare. The members of a gang will fight each other. They will even fight for a 'cause.' . . . Gangsters are impelled, in a way, to fight; so much of their activity is outside the law that fighting is the only means of avenging injuries and maintaining the code."

Thrasher mentioned a few gang-fights which went as far as homicide, but he did not give any comprehensive statistics of the quarrel-dead. I am indebted to Prof. Lundberg and Margaret Black Richardson, for the indication of Thrasher's work.

SUMMARY

The $\phi(\mu)$ diagram in Figure 1 is the best conspectus of the facts for the world as a whole. Alternative tracks are shown across the gap where world-totals are lacking. Samples from Manchoukuo and Chicago confirm the slope of the straight alternative, whereas continuity of slope is a consideration in favor of the circular arc.

MORE EXACT CONNECTION BETWEEN THE $\phi(\mu)$ GRAPH OF FIGURE 1 AND THE NUMBERS OF FATAL QUARRELS IN UNIT RANGES OF MAGNITUDE. INTERPOLATION ACROSS THE GAP.

We need to connect the mean over a range with the value at its midpoint. It might be thought that the connection was obvious; for a glance at Figure 1 shows that the graph is fairly straight; and of course the mean ordinate of a straight segment is simply the ordinate at its midpoint. But that would be a delusion based on a doubly wrong type of mean. For here the mean of q, not of the ordinate $\log_{10} q$, has to be taken over the integers in the range of n, and not over the abscissa $\mu = \log_{10} n$. It will be shown that the logarithmic transformation of both coordinates introduces correcting factors which depend on the slope, vanishing when it is either 0 or -2, but otherwise often considerable. They are obtained from the simplest formula that can fit the facts tolerably, namely the formula for a straight line on the $\phi(\mu)$ graph. The straight lines are suitably short segments; for it would be a mistake to smooth away significant detail by the over-wide sweep of any too-simple formula. In particular the sharp bend near the top of Figure 1 comes from British criminal statistics where it appears to be significant; and something rather like it is to be seen in the graphs which represent banditry in Manchoukuo and ganging in Chicago; so this bend should certainly not be smoothed away. But elsewhere a short length of the $\phi(\mu)$ graph can be represented by

$$\phi(\mu) = B - C\mu$$

where B and C are positive constants. On taking antilogarithms this becomes

$$q = 10^{B}n^{-c}$$

an inverse power law. To find $\bar{q}$ this n^{-c} has to be averaged over a range of n. As n proceeds only by integers, the average should be made by sums, not by integrals. However the integral $\int n^{-c}dn$ gives a useful approximation to the sum.

This programme will now be worked out in proper detail.

The notation is that already defined in connection with the graph. The typical range ends, say, at $\mu=\mu_0\pm\frac{1}{2}$, where $\mu_0=7, 6, 5, 4, 3, 2, 1$. It is convenient to begin with a rough and ready approximation, R, to the number Q of fatal quarrels in the range, whereby $\bar{q}(n)$ at the midpoint is multiplied by the number of integers in the range of quarrel-dead, which it is convenient for a later purpose, to denote by $\sum_n 1$ where $\sum_n$ sums for all n in the range of μ. Accordingly

$$R = 10^{\phi(\mu_0)}\sum_n 1. \tag{12}$$

For example, as read from the circular arc on Figure 1:

Range of magnitude $\mu=\log_{10}n$	Ordinate at midpoint $\phi(\mu_0)$	$\therefore \bar{q}(n)$ at midpoint	Number of integers in the range of quarrel-dead $\sum_n 1$	First approximation to number of quarrels R
0.5 to 1.5	3.87	7410	28	207,000
1.5 to 2.5	1.23	17.0	285	4,850
2.5 to 3.5	$\bar{1}$.17	0.131	2846	373

The above easy method is however suspect, because $\bar{q}(n)$ varies in the ratio 1740 in one of the ranges thus represented by its midpoint. So we must consider corrections.

Correction for width of range

As the $\phi(\mu)$ graph is nearly or quite straight, let us represent a local portion of it by

$$\phi(\mu) = \phi(\mu_0) - c(\mu - \mu_0) \tag{13}$$

where μ_0 is the midpoint of the range and $c=-d\phi/d\mu$ there. The distinction between $q(n)$ and $\bar{q}(n)$ is one that concerns detailed observations. It does not arise when, as now, deductions from a smooth curve in question. From (13)

$$q(n) = 10^{\phi(\mu)} = 10^{\phi(\mu_0)+c\mu_0} n^{-c}. \tag{14}$$

Let Q denote the total number of fatal quarrels in the range, then

$$Q = \sum_n q(n) = 10^{\phi(\mu_0)+c\mu_0} \sum_n n^{-c} \tag{15}$$

So from (12) and (15)

$$Q = R10^{c\mu_0} \sum_n n^{-c} \Big/ \sum_n 1. \tag{16}$$

The coefficient of R will be called the correcting factor. Two verifications of (16) may be noted: (i) if $c=0$, then $Q=R$. (ii) if the range of quarrel-dead contained only a single integer n, and if it were located at μ_0 so that $10^{\mu_0}=n$, then (16) would become $Q=Rn^c n^{-c}/1=R$, as it should.

The sum in (16) was evaluated by a method connected with the "integral test" for the convergence of series, in the following manner. Let

$$b = 10^{\mu_0+1/2} \quad \text{and} \quad a = 10^{\mu_0-1/2} \text{ for brevity.} \tag{17}$$

In the following chain of equations, the first is an approximation which improves as μ_0 increases, and the other are accurate.

$$\sum_n n^{-c} \simeq \int_a^b t^{-c}dt = \frac{1}{1-c}(b^{1-c} - a^{1-c})$$

$$= \frac{10^{\mu_0(1-c)}}{1-c}\left\{10^{(1-c)/2} - 10^{(c-1)/2}\right\}. \tag{18}$$

In the special case of $c=0$, (18) gives the approximation

$$\sum_n 1 = 10^{\mu_0}(10^{1/2} - 10^{-1/2}) = b - a. \tag{19}$$

Strictly $\sum_n 1$ is the integer next below $b-a$; but it is advisable to use the same type of approximation in both numerator and denominator of (16). Insertion of (18) and (19) into (16) gives the correcting factor

$$Q/R = \frac{1}{1-c} \cdot \frac{10^{(1-c)/2} - 10^{(c-1)/2}}{10^{1/2} - 10^{-1/2}}. \tag{20}$$

This expression is unity when $c=0$ or $c=2$.
For the wars, $c=1.50$ and $Q/R=0.855$.

The observed quantity is Q in the unit range of magnitude. The $\phi(\mu)$ curve should however be connected to μ at a point, and so to R. We have

$$\log_{10} R = \log_{10} Q - \log_{10} 0.855 = \log_{10} Q + 0.068.$$

In the $\phi(\mu)$ graph (Figure 1) the sloping line for the wars was therefore raised 0.068 above that which best fits the observations grouped in unit ranges of magnitude.

In the unexplored region we have, on taking the first approximation R from the previous table:

World-totals for the 126 years 1820 to 1945

Range of magnitude μ	$c = -\frac{d\phi}{d\mu}$	Correcting factor Q/R	Corrected number of fatal quarrels Q	Range of number killed in a quarrel n
0.5 to 1.5	3.14	1.92	397,000	4 to 31
1.5 to 2.5	2.32	1.16	5,630	32 to 316
2.5 to 3.5	1.87	0.95	354	317 to 3162

The fourth column shows the improved estimates for the unexplored region between the wars and the murders according to the circular arc in Figure (1). The difficulties of direct counting have already been described, and they are emphasized by these large numbers. In the range where 354 are expected, I have counted 188. If the straight line in Figure (1) were accepted instead of the circular arc, 354 would be increased to 2530. Although I know that the search is incomplete, I am unable to believe that less than a tenth have been found; and so I prefer the circular arc.

THE TOTAL NUMBER OF PERSONS WHO DIED BECAUSE OF QUARRELS DURING THE 126 YEARS FROM 1820 TO 1945 A.D.

Ends of range of magnitude	Total number of deaths in millions	How computed
$7 \pm \frac{1}{2}$	36	By summation over a list of fatal quarrels.
$6 \pm \frac{1}{2}$	6.7	
$5 \pm \frac{1}{2}$	3.4	
$4 \pm \frac{1}{2}$	0.75	
$3 \pm \frac{1}{2}$	0.30	From the circular arc in Figure (1) by the method described below.
$2 \pm \frac{1}{2}$	0.40	
$1 \pm \frac{1}{2}$	2.2	
$0 \pm \frac{1}{2}$	9.7	From page 528
	Total 59	

A remarkable feature of the above table is that the heavy loss of life occurred at the two ends of the sequence of magnitudes, namely the World Wars and the murders. The small wars contributed much less to the total. The total deaths because of quarrels should be compared with the *total deaths from all causes*. There are particulars given by de Jastrzebski [9], and in the Statistical Year Books of the League of Nations, which allow the total to be estimated. A mean world population of 1.5×10^9 and a mean death rate of 20 per thousand per year would give during 126 years 3.8×10^9 deaths from all causes. Of these the part caused by quarrels was 1.6 per cent. This is less than one might have guessed from the large amount of attention which quarrels attract. Those who enjoy wars can excuse their taste by saying that wars after all are much less deadly than disease.

The method of computation in the unexplored region was an extension of that explained above. The number D of dead in any range of μ is strictly

$$D = \sum_n nq(n) = \sum_n 10^{\mu+\phi(\mu)}. \tag{21}$$

First a common-sense estimate, E of D was obtained from the graph at the midpoint μ_0 of the range $\mu_0-\frac{1}{2}\leqq\mu\leqq\mu_0+\frac{1}{2}$ thus, by (12),

$$E = 10^{\mu_0}R = 10^{\mu_0+\phi(\mu_0)}\sum_n 1. \tag{22}$$

Then E was corrected by multiplication by a factor which was found to be approximately

$$\frac{D}{E} = \frac{1}{2-c}\left\{\frac{10^{(2-c)/2} - 10^{(c-2)/2}}{10^{1/2} - 10^{-1/2}}\right\}. \tag{23}$$

The theory of this correction is based on the approximation (13), whence it follows that

$$\mu + \phi(\mu) = c\mu_0 + \phi_0 + \mu(1-c). \tag{24}$$

From (21) and (24)

$$D = 10^{c\mu_0+\phi(\mu_0)}\sum_n n^{1-c}, \quad \text{and} \tag{25}$$

$$\frac{D}{E} = 10^{\mu_0(c-1)}\,\frac{\sum_n n^{1-c}}{\sum_n 1}. \tag{26}$$

The sums in (26) were obtained from (18) by suitable alterations of the index and they lead to (22). The correction (22) was verified in particular cases by the "deferred approach to the limit" of Richardson and Gaunt [14].

COMPARISON WITH THE DISTRIBUTION OF SIZES OF TOWNS

Towns and wars are both examples of human aggregation. A referee has asked me to compare them. For simplicity I take the distribution of towns from Lotka's book [10A] in Auerbach's idealized form whereby the town of rank r in a given country has a population n such that

$$nr = A, \tag{1}$$

in which A is a constant, namely the population of the largest town. The desired comparison might conceivably be made by arranging the fatal quarrels in order of rank; but there would be many artificially coincident ranks on account of the rounding off of imperfectly known casualties; moreover the gap in the data prevents the extension of rank from the wars to the murders. It is preferable therefore to leave the fatal quarrels as they are already shown on Figure 1, and to transform Auerbach's law so as to see how towns would appear on a diagram of that type. For the purpose of comparison the same symbols will be used in corresponding meanings. Thus in this section

$$n \text{ is the population of a town, and } \mu \text{ will be } \log_{10} n. \tag{2}$$

Again, to match the fatal quarrels, $\bar{q}$ will here be defined as

$$\bar{q} = \frac{\text{number of towns in a range of population}}{\text{number of integers in that range of population}}. \tag{3}$$

Also, as before, let

$$\phi = \log_{10} \bar{q}. \tag{4}$$

Consider two towns of ranks r', r'', of which r'' is the greater. Let them mark the ends of a range. The number of towns inside the range is strictly $r''-r'-1$. But let us follow the usual statistical practice of regarding half of an end-object as lying on either side of the end. The number of towns in the range is then simply $r''-r'$. Let n' and n'' be

the respective populations. With the same convention about ends on the population range, we have from (3)

$$\bar{q} = \frac{r'' - r'}{n' - n''} \cdot \tag{5}$$

Elimination of ranks between (5) and (1) gives

$$\bar{q} = \frac{A}{n'n''} \cdot \tag{6}$$

So that from (4)

$$\phi = \log_{10} A - \log_{10} n' - \log_{10} n''. \tag{7}$$

As for the wars, let μ_0 be the midpoint of the range of μ. It is

$$\mu_0 = \tfrac{1}{2}(\log_{10} n' + \log_{10} n''). \tag{8}$$

From (7) and (8)

$$\phi = \log_{10} A - 2\mu_0, \tag{9}$$

however wide or narrow the range may be. That the constant $\log_{10} A$ should be the same whatever the grouping is a peculiarity of the slope

$$d\phi/d\mu_0 = -2 \tag{10}$$

as shown on page 541, where the correction vanished for $c=2$.

This slope for the towns is quite close to that of the straight line which in Figure 1 might join the murders to the world wars. Such a broad resemblance between two forms of aggregation is certainly interesting, and may suggest theories. Perhaps the most suggestive formula is

$$q = An^{-2} \tag{11}$$

which follows from (6) when n' is almost equal to n''. Rashevsky [11A], who gives formula (11), has tried-out several explanations.

These overall resemblances do not however conduce to an accurate description of fatal quarrels; for they distract attention from the curvature of the $\phi(\mu)$ graph in Figure 1. The slope of the part relating to wars is certainly not -2. If it were so, then the number of fatal quarrels would increase 10 times for each unit decrease of magnitude. The observed ratio is less than four.

REFERENCES

[1] Anderson, J. G. 1939. *China fights for the world.* Kegan Paul, London, p. 106.
[1A] Carr-Saunders, A. M. 1936. *World Population.* Clarendon Press, Oxford.
[2] Coupland, R. 1939. *The Exploitation of East Africa 1856–1890.* Faber & Faber, London. Chh. VII, XI.
[3] Durbin, E. F. M., and Bowlby, J. 1939. *Personal Aggressiveness and War.* Kegan Paul, London.
[4] Fisher, R. A. 1922. *Phil. Trans. Roy. Soc. Lond.* A **222**, 332.
[5] Glover, E. 1933 & 1947. *War, Sadism & Pacifism.* Allen & Unwin, London.
[6] Harding, D. W. 1941. *The Impulse to Dominate.* Allen & Unwin, London.
[7] Hume, M. A. S. 1900. *Modern Spain.* Fisher Unwin, London.
[8] *Japan and Manchoukuo Year Book,* 1938. Tokio.
[9] Jastrzebski, S. de. 1929 article *Death-Rate* in Ency. Brit. XIV edn.
[10] Jeffreys, H. 1939. *Theory of Probability,* p. 148. Clarendon Press, Oxford.
[10A] Lotka, A. J. 1925. *Elements of Physical Biology.* Williams & Wilkins, Baltimore.
[11] Miliukov, P. article *Russia* in *Ency. Brit.* XIV ed., **19**, 720 d.
[11A] Rashevsky, N. 1947. *Mathematical Theory of Human Relations.* The Principia Press, Bloomington, Indiana.
[12] Richardson, L. F. in *Nature* of 15 Nov. 1941.
[13] Richardson, L. F. 1945. *J. Roy. Statistical Soc.* **107**, 242–250.
[14] Richardson, L. F., and Gaunt, J. A. 1927. *Phil. Trans. Roy. Soc. Lond.* A **226**, 299–361.
[15] Sleeman, W. H. 1836. *Ramaseeana.* Military Orphan Press Calcutta, p. 39.
[16] Thrasher, F. M. 1927. *The Gang.* Chicago University Press.
[17] Whitaker's Almanack, annual, London.
[18] Wright, Q. 1942. *A Study of War.* Chicago University Press.

BANK RESERVES AND BUSINESS FLUCTUATIONS*

CLARK WARBURTON
Federal Deposit Insurance Corporation

The belief that severe business fluctuations are caused primarily by the operations of banks, through their influence on the quantity of the circulating medium, was widely held during the nineteenth century and the first three decades of the twentieth. During the latter part of this period variations in bank reserves were believed to be the chief causal factor underneath changes in the circulating medium furnished by the banks. This paper describes a study of factual data since 1918 in the light of that theory. To make the investigation, it was necessary to develop several new time series from available statistical material, and to make appropriate allowance for growth trends. The results support the theory. There is a case by case correspondence between the peaks and troughs in the series of effective member bank reserves and the business cycle reference dates of the National Bureau of Economic Research, with leadership in reserves; and the duration and depth of periods of recession and recovery intervening between the turning points are closely associated with the duration and depth of deviations in reserves from the line of growth.

I WOULD like to begin with two quotations. The first is from Harry E. Miller's book, *Banking Theories in the United States before 1860.*

> "The most common explanation of commercial crises . . . was that they result from the alternating expansion and contraction of the media of payments which occurs wherever a currency of convertible bank notes exists."[1]

The second is from Paul Barnett's book, *Business-Cycle Theory in the United States from 1860 to 1900.*

> "Monetary theories of cyclical changes in business activity received more widespread consideration by writers in the United States from 1860 to 1900 than any other single group of theories."[2]

---

* Talk before the Washington Statistical Society, April 26, 1948.

[1] Harry E. Miller, *Banking Theories in the United States before 1860* (Harvard University Press, 1927), p. 81.

[2] Paul Barnett, *Business-Cycle Theory in the United States, 1860–1900* (University of Chicago Press, 1941), p. 86.

The belief that severe business fluctuations are caused primarily by the operations of banks, through their influence on the quantity of the circulating medium, continued to be widely held during the first three decades of the present century. During this period and the latter part of the nineteenth century an important supplement to the theory was developed, namely, that variations in bank reserves are the chief causal factor underneath changes in the circulating medium furnished by the banks. In addition, the need for continuous increase in the quantity of circulating medium because of increasing population and productivity was given emphasis; the possibility that habits of use of circulating medium may change was recognized; and various aspects of the routes or modes of impact of bank operations were explored.

To describe with precision all phases of this banking theory of business fluctuations cannot be attempted here. My endeavor to provide a concise but reasonably complete statement of the theory has yielded the following six points.

1. Under a fractional reserve system, changes in the amount of commercial bank reserves alter the quantity of deposits and therefore the quantity of assets which such banks may hold; and this is followed by changes in the actual volume of assets and deposits, because banks of deposit have an inherent tendency to expand their operations to the limit placed upon them by law and interbank relationships. Since deposits form the bulk of the circulating medium, changes in deposits dominate changes in the total circulating medium.

2. Changes in the amount of circulating medium, expressed as deviations from a normal rate of increase accompanying growth in population and productivity and adjusted for any measurable trend in monetary habits, alter the equilibrium center of the price system. That is, deficiency in the circulating medium relative to a reasonable rate of growth produces a falling price level because the output of the economy cannot continue to be sold at prevailing prices; while expansion at an excessive rate gives sellers of goods an opportunity to obtain higher prices and the desire for profits produces this result.

3. A falling price level, in view of fixed and sticky costs, narrows profit margins and induces business to reduce output; while a rising price level tends to increase profit margins and therefore stimulates output. In addition, reluctance or inability of some business enterprises to make rapid changes in selling prices leads to unanticipated inventory accumulation when the equilibrium center of prices falls and to diminished inventories when the equilibrium center rises. Such changes in

inventories, as well as changes in prices, induce businessmen to curtail or expand output.

4. The foregoing effects are likely to be accentuated by changes in velocity of the circulating medium. Contraction of the circulating medium is *per se* a shrinking of the cash holdings of business and individuals, and this leads to more hesitant and cautious use of them. On the other hand, rapid expansion of cash balances is followed by a tendency to shorten the period that cash receipts are held before their disbursement.

5. An additional route of impact of changes in bank reserves is through accompanying changes in rates of interest and in the availability of credit for those who wish to borrow. Since those changes may be synchronous with the beginning of a deviation of bank reserves from trend, the effects of such deviations via this route of impact may precede those via the other routes which I have described.

6. Causation sequences are normally accompanied by time lags. Two such lags between changes in bank reserves and sequential business fluctuations are important: (a) that which may occur between the appearance of the forces operating to change the quantity of reserves and the response of commercial banks to those changes; and (b) that which may occur between changes in the quantity of circulating medium and changes in the level of prices and quantity of output. These time lags vary in length. Under some circumstances the total lag may be of negligible duration; under other circumstances as much as or more than a year.

This hypothesis, as I have indicated, was the dominant aspect of business-fluctuation theory in the nineteenth and early part of the twentieth centuries. However, it was far from universally held, and the validity of various parts of it was challenged by some economists. Efforts at testing it by factual analysis were inconclusive because of inadequacy of data. In recent years the hypothesis has been displaced by other theories to such a great extent that few economists have deemed it worthwhile to examine in its light the more abundant data now available. My purpose in offering this paper is to describe briefly the statistical methodology which I have used in such an examination and to state the results.

Scope of statistical analysis

Variations in the value, quantity, and prices of goods and services sold to final purchasers, in the form of deviations from appropriate

trends, were deemed to constitute the basic phenomena of business fluctuations. These, together with the quantity of circulating medium (measured by total deposits and currency excluding interbank obligations), habits of use of circulating medium as expressed in velocity, and bank reserves were selected as the principal statistical series to be examined.

The period from 1918 to date was chosen for the study for two reasons. First, since June 1917, required reserves of banks which are members of the Federal Reserve System have consisted exclusively of deposit balances in the Federal Reserve Banks. Second, most of the necessary statistical series were unavailable for prior dates, nor could reasonably reliable series be constructed from the available information.

Quarterly averages were used in the study. On the one hand annual data are not sufficiently sensitive to the variations commonly designated "cyclical"; and on the other, monthly data are more difficult to work with over a period of three decades and are also subject to short-time variations which may be neglected.

Statistical series

Existing statistics did not provide suitable basic series and these had to be constructed. With the exception of bank reserves, the methods of constructing them are described in articles which have been published in economic journals.[3] These series have since been revised, with slight changes in methodology. With respect to bank reserves, reported data for Federal Reserve member banks were adjusted to quarterly averages, and then adjusted successively for changes in percentage reserve requirements, for shifting of deposits among member banks subject to different percentage requirements, and for changes in the proportion of deposits held by member and nonmember banks. These adjustments, together with the resulting series of effective bank reserves, are shown in Table 3, at the end of this article.

Determination of trends

For initial determination of trends in sales of final products and in the quantity and rate of use of circulating medium, the period of greatest stability of output and prices during the three decades was selected.

[3] "Normal Production, Income, and Employment, 1945 to 1965," *The Southern Economic Journal*, Vol. XI, No. 3 (January 1945), pp. 219–45; "The Volume of Money and the Price Level Between the World Wars," *The Journal of Political Economy*, Vol. LIII, No. 2 (June 1945), pp. 150–63; and "Quantity and Frequency of Use of Money in the United States, 1919–45," *The Journal of Political Economy*, Vol. LIV, No. 5 (October 1946), pp. 436–50.

At first the seven-year period, 1923–1929, was used, but in view of deviations which occurred in 1929 the period was shortened to the six years, 1923–1928. Percentage growth trends were used, based initially on averages for the terminal years of the period, 1923 and 1928, both years of high production or "full employment" as that phrase is now used.

The trends thus derived were drawn on semi-logarithmic charts of the data for the entire period covered by the study, and were found to be excellent fits for this period. In addition, the trend in quantity of final products sold was compared with an annual index of production covering the period since 1879; and the trend in circuit velocity of deposits and currency was compared with the ratio of national income to deposits and currency by decades for the period since 1799. The apparent trends for the recent and the earlier decades were similar, and their validity as genuine secular trends was deemed to be established.

Several minor variations in the slopes of the trends were examined to see which provided the best theoretical and practical fit. The final selection took into account the fact that the level of prices of final products was lower in 1928 than in 1923 by about 2 per cent, and the principle that the trend values should be consistent with each other when projected forward for a considerable period of time. The slopes chosen were as follows:

Prices of final products: none.
Quantity and value of final products: 3.6 per cent growth per year, or .888 per cent per quarter.
Circuit velocity of deposits and currency: inverse of a growth in deposits and currency relative to expenditures for final products of 1.35 per cent per year, or .336 per cent per quarter.
Quantity of deposits and currency: 5 per cent growth per year, or 1.227 per cent per quarter (when expressed as amounts, the product of the above two trends).
Bank reserves: same as quantity of deposits and currency.

In the final revision, the selected quarterly trend rates were applied to the 25-quarter period from the last quarter of 1922 to the last quarter of 1928, inclusive. Trend values were computed from the quarter in the center of the period.

Deviations from trend

The quarterly statistical series were next compared with trend values.

Points of upward and downward departure from trend, and of maximum deviation, were noted; and also differences in timing with respect to turning points and departures from trend in the various series. In the case of the transition from the prosperity of the 1920's to the great depression, the timing of departures from trend, on the down side, were as follows:

Bank reserves, and deposits and currency—third quarter of 1928
Value of sales of final products—second quarter of 1929
Quantity of sales of final products—fourth quarter of 1929
Prices of final products—first quarter of 1930
Circuit velocity of deposits and currency—third quarter of 1930

This sequence was found to be typical, though not perfectly uniform, and with considerable variation in the length of the lags.

One of the important results is the fact that deviations from trend in circuit velocity—particularly in the case of the beginnings of downswing—lag behind deviations from trend in quantity of deposits and currency, and except for this lag are in the same direction. This result is in conformity with traditional theory that variation in velocity results from and accentuates the effect of variation in the quantity of circulating medium. The assumption which is made in Keynesian and other contemporary theories that idle money is an independent or leading factor in the course of business fluctuations is inconsistent with the observed facts.

Comparison of turning points with business cycle reference dates

The last element in the statistical methodology to be mentioned here was a comparison, for the period 1918–1939, of turning points in bank reserves and in deposits and currency with each other and with reference dates for business cycles. For turning points in bank reserves and in deposits and currency, downswings and upswings relative to trend of one quarter's duration only were neglected. Other turning points were divided into two groups. Those preceded and succeeded by an upswing or downswing of a year or longer were defined as primary turning points; those preceded or succeeded by a downswing or upswing of two or three quarter's duration were termed secondary turning points. The secondary turning points occur in pairs between a pair of primary turning points.

For the period from 1918 to 1939 five pairs of primary turning points

and six pairs of secondary turning points occurred in bank reserves. Of the latter, four pairs showed slight amplitude, that is, the difference between the peak and trough expressed as deviations from trend is about one per cent or less; the other two were of considerably greater severity.

TABLE I

TURNING POINTS IN MEMBER BANK RESERVES AND IN DEPOSITS AND CURRENCY COMPARED WITH BUSINESS CYCLE REFERENCE DATES, QUARTERLY DATA, 1918–1939

Turning points	Effective member bank reserves*	Deposits and currency†	Business cycle reference dates‡
Peak	2nd 1918	..	3rd 1918
Trough	4th 1918	..	2nd 1919
Peak	1st 1920	2nd 1920	1st 1920
Trough	4th 1921	1st 1922	3rd 1921
Peak	1st 1923	2nd 1923	2nd 1923
Trough	2nd 1924	2nd 1924	3rd 1924
Peak	4th 1925	4th 1925	3rd 1926
Trough	1st 1927	1st 1927	4th 1927
Peak	2nd 1928	2nd 1928	2nd 1929
Trough	1st 1932	..	Summer 1932
Peak	4th 1932	..	End of 1932
Trough	2nd 1933	2nd 1933	1st 1933
Peak	4th 1935	3rd 1936	2nd 1937
Trough	3rd 1937	2nd 1938	2nd 1938

* Based on deviations from trend, at 5 per cent per year rate of growth, in average member bank reserves adjusted for changes in percentage reserve requirements, shifting of deposits among member banks subject to different percentage requirements, and changes in the proportion of deposits held by member and nonmember banks. Includes all peaks and troughs in this series preceded by upswings and downswings, respectively, of a year or longer, and all peaks and troughs preceded or followed by an upswing or downswing of two or three quarters' duration with an amplitude of 2 per cent or more.

† Based on deviations from trend in total deposits and currency, excluding interbank obligations—with the same criteria as for member bank reserves. In the omitted cases the quarterly estimates of deposits and currency do not show definite peaks and troughs, though changes occurred in the rate of expansion or contraction relative to trend.

‡ Arthur F. Burns and Wesley C. Mitchell, *Measuring Business Cycles* (1946), 78 and 82. The dates given here include all of the reference cycles recognized by Burns and Mitchell for the period subsequent to January 1, 1918, and also the intermediate cycle of short duration but large amplitude in 1932, which they mention but treat as a double-bottomed trough without assignment of specific reference dates.

The list of turning points in member bank reserves given in Table 1 includes all primary turning points and the two pairs of secondary turning points with considerable amplitude. The business cycle reference dates include all those given for the period since 1918 by Professors Burns and Mitchell of the National Bureau of Economic Research, together with the intermediate cycle of short duration and substantial

amplitude in 1932 which they mention but treat as a double-bottomed trough.[4]

It will be noted that there is a case-by-case correspondence between the 14 turning points in member bank reserves during the period from 1917 to 1939 and the same number of National Bureau business cycle reference dates. The peaks and troughs in member bank reserves preceded the business cycle peaks and troughs in ten of the fourteen cases, were simultaneous in two cases, and lagged by one quarter in two cases. However, in the last two cases the specific data are such as to make unwarranted a conclusion that in those cases the recovery in business activity preceded the recovery in reserves.

TABLE 2

RELATIONSHIP OF DEGREE OF CHANGE IN THE EFFECTIVE AMOUNT OF MEMBER BANK RESERVES TO PERIODS OF BUSINESS RECOVERY AND RECESSION, 1918–1941*

Approximate period of expansion or contraction in reserves	Approximate period of recession or recovery
1918–1919, rapid expansion in reserves	1919 and early 1920, business boom
1920–1921, contraction in reserves	Late 1920 and 1921, severe depression
1922, rapid expansion in reserves	1922 and 1923, prosperity
1923, insufficient growth in reserves	1924, minor depression
1924, rapid expansion in reserves	1925, boom
1925, normal growth in reserves	1926, continuation of prosperity
1926, contraction in reserves	1927, depression
1927, rapid expansion in reserves	1928 and part of 1929, prosperity
1928–1930, insufficient growth in reserves	Late 1929 and 1930, depression
1931 and early 1932, drastic contraction in reserves	1932, intensified depression
Summer and autumn of 1932, substantial expansion in reserves	Latter part of 1932, beginning of recovery
Early 1933, drastic contraction in reserves	Same period, renewed depression and collapse of the banking system
Middle 1933 to end of 1935, rapid growth in reserves	Mid-1933 to 1937, continuous business improvement
1936–1937, drastic contraction in reserves	Late 1937 and early 1938, depression
1938–1940, rapid expansion in reserves	Late 1938 to Pearl Harbor, prosperity.

* For description of the concept of effective member bank reserves, see footnote * to table 1.

The relationship of bank reserves to business fluctuations is not limited to the sequence of turning points; the duration and depth of the periods of recovery and recession intervening between the turning points are closely associated with the duration and depth in deviations in effective member bank reserves from the 5 per cent per year line of growth. The sequences are shown in Table 2.

[4] Arthur F. Burns and Wesley C. Mitchell, *Measuring Business Cycles* (National Bureau of Economic Research, 1946), pp. 78 and 82.

TABLE 3

MEMBER BANK RESERVES, ANNUAL AND QUARTERLY AVERAGES, 1918-1947

Year and quarter	Amount in millions of dollars				Relatives for computation of effective amount (1923-1928=1)			Percentage ratio to trend of effective amount of reserves[7]
	Actual amount			Effective amount (1923-28 basis)[3]	Percentage requirement[4]	Deposit shifting[5]	Deposits not requiring reserve and interbank deposits[6]	
	Total[1]	Required[2]	Excess					
1918	*1,496*	*1,382*	*114*	*1,437*	*1.00*	*1.158*	*1.112*	*96.8*
I	1,467	1,377	90	1,411	1.00	1.163	1.119	96.3
II	1,499	1,298	201	1,480	1.00	1.156	1.141	99.9
III	1,471	1,319	152	1,432	1.00	1.154	1.123	95.4
IV	1,548	1,535	13	1,424	1.00	1.158	1.065	93.7
1919	*1,718*	*1,636*	*82*	*1,600*	*1.00*	*1.158*	*1.059*	*102.1*
I	1,633	1,559	74	1,479	1.00	1.154	1.045	96.2
II	1,679	1,607	72	1,561	1.00	1.145	1.065	100.3
III	1,743	1,652	91	1,644	1.00	1.133	1.069	104.3
IV	1,817	1,726	91	1,717	1.00	1.119	1.057	107.6
1920	*1,835*	*1,772*	*63*	*1,745*	*1.00*	*1.092*	*1.037*	*106.0*
I	1,873	1,771	102	1,766	1.00	1.107	1.044	109.3
II	1,859	1,804	55	1,749	1.00	1.097	1.032	107.0
III	1,821	1,808	13	1,727	1.00	1.087	1.031	104.4
IV	1,785	1,705	80	1,730	1.00	1.075	1.042	103.3
1921	*1,672*	*1,623*	*49*	*1,656*	*1.00*	*1.060*	*1.049*	*95.8*
I	1,732	1,642	90	1,709	1.00	1.064	1.050	100.8
II	1,662	1,629	33	1,657	1.00	1.057	1.054	96.5
III	1,630	1,586	44	1,620	1.00	1.058	1.051	93.2
IV	1,663	1,636	27	1,636	1.00	1.059	1.042	93.0
1922	*1,781*	*1,738*	*43*	*1,718*	*1.00*	*1.052*	*1.016*	*94.7*
I	1,702	1,641	61	1,659	1.00	1.056	1.029	93.1
II	1,779	1,739	40	1,710	1.00	1.054	1.013	94.9
III	1,807	1,785	22	1,730	1.00	1.053	1.008	94.8
IV	1,834	1,787	47	1,774	1.00	1.046	1.012	96.0
1923	*1,873*	*1,824*	*49*	*1,873*	*1.00*	*1.019*	*1.019*	*98.3*
I	1,897	1,821	76	1,871	1.00	1.031	1.017	100.1
II	1,870	1,833	37	1,869	1.00	1.021	1.020	98.7
III	1,850	1,820	30	1,862	1.00	1.014	1.021	97.2
IV	1,874	1,822	52	1,889	1.00	1.011	1.019	97.4
1924	*2,023*	*1,962*	*61*	*1,989*	*1.00*	*1.019*	*1.003*	*99.5*
I	1,906	1,837	69	1,913	1.00	1.011	1.015	97.5
II	1,943	1,902	41	1,929	1.00	1.015	1.008	97.1
III	2,079	1,997	82	2,030	1.00	1.023	.999	100.9
IV	2,162	2,112	50	2,082	1.00	1.027	.989	102.3
1925	*2,166*	*2,130*	*36*	*2,134*	*1.00*	*1.010*	*.995*	*101.7*
I	2,163	2,104	59	2,101	1.00	1.016	.987	101.9
II	2,132	2,102	30	2,109	1.00	1.007	.996	101.1
III	2,157	2,115	42	2,140	1.00	1.007	.999	101.3
IV	2,214	2,200	14	2,187	1.00	1.009	.997	102.3
1926	*2,209*	*2,191*	*18*	*2,212*	*1.00*	*.999*	*1.000*	*100.3*
I	2,214	2,197	17	2,201	1.00	1.004	.998	101.7
II	2,196	2,172	24	2,207	1.00	.996	1.001	100.7
III	2,208	2,184	24	2,216	1.00	.998	1.002	99.9
IV	2,217	2,210	7	2,222	1.00	.997	.999	99.0

TABLE 3 (*CONTINUED*)

Year and quarter	Amount in millions of dollars				Relatives for computation of effective amount (1923–1928 = 1)			Percentage ratio to trend of effective amount of reserves[7]
	Actual amount			Effective amount (1923–28 basis)[3]	Percentage requirement[4]	Deposit shifting[5]	Deposits not requiring reserve and interbank deposits[6]	
	Total[1]	Required[2]	Excess					
1927	*2,290*	*2,286*	*4*	*2,298*	*1.00*	*.987*	*.991*	*99.3*
i	2,232	2,210	22	2,245	1.00	.990	.996	98.8
ii	2,270	2,266	4	2,284	1.00	.987	.993	99.3
iii	2,291	2,285	6	2,300	1.00	.985	.989	98.8
iv	2,366	2,381	−15	2,362	1.00	.986	.984	100.2
1928	*2,355*	*2,342*	*13*	*2,415*	*1.00*	*.973*	*.997*	*99.4*
i	2,386	2,381	5	2,393	1.00	.984	.987	100.3
ii	2,380	2,368	12	2,444	1.00	.971	.997	101.2
iii	2,304	2,292	12	2,393	1.00	.965	1.002	97.9
iv	2,350	2,362	24	2,428	1.00	.970	1.002	98.1
1929	*2,358*	*2,315*	*43*	*2,472*	*1.00*	*.967*	*1.014*	*96.9*
i	2,360	2,314	47	2,473	1.00	.962	1.008	98.7
ii	2,306	2,270	37	2,452	1.00	.960	1.021	96.7
iii	2,330	2,293	37	2,465	1.00	.964	1.020	96.0
iv	2,434	2,382	52	2,498	1.00	.981	1.007	96.1
1930	*2,378*	*2,323*	*55*	*2,463*	*1.00*	*.965*	*.999*	*91.9*
i	2,328	2,277	51	2,417	1.00	.963	1.000	91.9
ii	2,366	2,319	47	2,456	1.00	.964	1.001	92.2
iii	2,402	2,340	62	2,492	1.00	.963	.999	92.5
iv	2,418	2,357	61	2,488	1.00	.968	.996	91.2
1931	*2,323*	*2,235*	*89*	*2,396*	*1.00*	*.974*	*1.005*	*85.1*
i	2,396	2,321	76	2,448	1.00	.974	.995	88.6
ii	2,389	2,305	84	2,461	1.00	.968	.997	88.0
iii	2,362	2,247	115	2,443	1.00	.973	1.006	86.3
iv	2,148	2,066	82	2,233	1.00	.982	1.021	77.9
1932	*2,114*	*1,858*	*256*	*2,216*	*1.00*	*.984*	*1.033*	*75.0*
i	1,928	1,883	46	2,040	1.00	.976	1.033	70.3
ii	2,065	1,844	221	2,193	1.00	.978	1.039	74.7
iii	2,086	1,812	273	2,203	1.00	.981	1.036	74.1
iv	2,373	1,892	481	2,428	1.00	1.000	1.023	80.7
1933	*2,306*	*1,784*	*522*	*2,287*	*1.00*	*1.029*	*1.021*	*73.7*
i	2,218	1,817	401	2,240	1.00	1.009	1.019	73.6
ii	2,090	1,736	354	2,054	1.00	1.041	1.023	66.6
iii	2,334	1,776	559	2,314	1.00	1.032	1.023	74.2
iv	2,581	1,808	773	2,541	1.00	1.033	1.017	80.4
1934	*3,664*	*2,104*	*1,560*	*3,420*	*1.00*	*1.059*	*.991*	*105.0*
i	2,961	1,918	1,044	2,867	1.00	1.041	1.008	89.7
ii	3,689	2,072	1,616	3,476	1.00	1.056	.995	107.4
iii	3,973	2,164	1,809	3,671	1.00	1.065	.984	112.1
iv	4,034	2,263	1,771	3,666	1.00	1.073	.975	110.6
1935	*4,999*	*2,532*	*2,467*	*4,255*	*1.00*	*1.103*	*.941*	*124.5*
i	4,469	2,358	2,112	3,965	1.00	1.082	.960	118.1
ii	4,731	2,478	2,254	4,072	1.00	1.091	.939	119.8
iii	5,148	2,599	2,550	4,305	1.00	1.112	.930	125.1
iv	5,647	2,693	2,955	4,679	1.00	1.126	.933	134.4

TABLE 3 (*CONTINUED*)

Year and quarter	Amount in millions of dollars					Relatives for computation of effective amount (1923–1928=1)			Percentage ratio to trend of effective amount of reserves[7]	
	Actual amount			Effective amount (1923–28 basis)[2]		Percentage requirement[4]	Deposit shifting[5]	Deposits not requiring reserve and interbank deposits[6]		
	Total[1]	Required[2]	Excess							
1936	*5,988*	*3,476*	*2,512*	*4,008*		*1.25*	*1.135*	*.923*	*111.6*	
i	5,669	2,761	2,908	4,666		1.00	1.131	.931	132.4	
ii	5,474	2,840	2,634	4,477		1.00	1.131	.925	125.5	
iii	6,129	3,723	2,406	3,307		1.50	1.137	.920	91.6	
iv	6,681	4,579	2,103	3,583		1.50	1.140	.917	98.0	
1937	*6,829*	*5,603*	*1,226*	*2,910*		*1.94*	*1.119*	*.921*	*77.2*	
i	6,722	4,850	1,872	3,103		1.75	1.134	.916	83.8	
ii	6,878	5,759	1,118	2,817		2.00	1.122	.919	75.2	
iii	6,800	5,958	842	2,820		2.00	1.113	.923	74.3	
iv	6,917	5,845	1,073	2,899		2.00	1.106	.927	75.5	
1938	*7,931*	*5,415*	*2,516*	*3,631*		*1.81*	*1.120*	*.925*	*91.7*	
i	7,246	5,818	1,428	3,043		2.00	1.106	.929	78.3	
ii	7,645	5,192	2,453	3,650		1.75	1.112	.929	92.8	
iii	8,161	5,194	2,967	3,837		1.75	1.123	.924	96.4	
iv	8,673	5,458	3,215	3,993		1.75	1.137	.916	99.1	
1939	*10,344*	*5,957*	*4,387*	*4,571*		*1.75*	*1.159*	*.898*	*100.0*	
i	8,992	5,562	3,430	4,101		1.75	1.139	.909	100.5	
ii	9,902	5,774	4,128	4,449		1.75	1.151	.905	107.7	
iii	10,808	6,072	4,736	4,735		1.75	1.166	.894	113.3	
iv	11,674	6,421	5,253	4,999		1.75	1.181	.885	118.1	
1940	*13,248*	*6,922*	*6,324*	*5,483*		*1.75*	*1.197*	*.868*	*125.6*	
i	12,187	6,579	5,608	5,141		1.75	1.187	.876	120.0	
ii	13,128	6,799	6,329	5,468		1.75	1.195	.871	126.1	
iii	13,595	7,015	6,580	5,595		1.75	1.201	.865	127.4	
iv	14,074	7,294	6,780	5,729		1.75	1.206	.859	128.9	
1941	*13,407*	*8,078*	*5,329*	*5,206*		*1.81*	*1.212*	*.849*	*113.6*	
i	14,105	7,586	6,519	5,679		1.75	1.212	.854	126.2	
ii	13,483	7,796	5,688	5,394		1.75	1.214	.850	118.4	
iii	13,103	7,994	5,109	5,224		1.75	1.214	.847	113.3	
iv	12,936	8,935	4,001	4,525		2.00	1.208	.845	97.0	
1942	*12,649*	*9,976*	*2,673*	*4,413*		*2.00*	*1.187*	*.836*	*91.7*	
i	12,890	9,590	3,301	4,456		2.00	1.215	.840	94.3	
ii	12,594	9,320	2,774	4,293		2.00	1.219	.831	89.8	
iii	12,444	10,182	2,262	4,185		2.00	1.222	.822	86.5	
iv	12,668	10,313	2,355	4,719		2.00	1.090	.812	96.3	
				A	B			A B	A	B
1943	*12,628*	*11,117*	*1,511*	*5,107*	*4,793*	*2.00*	*1.080*	*.874 .820*	*101.1*	*94.9*
i	13,177	11,268	1,909	4,982	4,982	2.00	1.095	.828 .828	100.4	100.4
ii	12,345	10,631	1,713	4,965	4,656	2.00	1.079	.868 .814	98.9	92.7
iii	12,624	11,330	1,293	5,240	4,783	2.00	1.077	.894 .816	103.1	94.1
iv	12,366	11,238	1,128	5,239	4,752	2.00	1.067	.904 .820	101.8	92.4
1944	*13,220*	*12,170*	*1,050*	*6,107*	*5,199*	*2.00*	*1.055*	*.973 .829*	*115.1*	*98.0*
i	12,572	11,484	1,087	5,516	4,918	2.00	1.061	.931 .830	105.9	94.4
ii	13,018	12,069	947	5,981	5,036	2.00	1.060	.974 .820	113.4	95.5
iii	13,127	12,051	1,074	6,226	5,206	2.00	1.049	.995 .832	116.7	97.6
iv	14,165	13,075	1,088	6,704	5,636	2.00	1.048	.992 .834	124.1	104.3

TABLE 3 (*CONTINUED*)

Year and quarter	Amount in millions of dollars					Relatives for computation of effective amount (1923–1928 = 1)				Percentage ratio to trend of effective amount of reserves[7]	
	Actual amount			Effective amount (1923–28 basis)[3]		Percentage requirement[4]	Deposit shifting[5]	Deposits not requiring reserve and interbank deposits[6]			
	Total[1]	Required[2]	Excess	A	B			A	B	A	B
1945	*15,050*	*13,928*	*1,122*	*7,340*	*6,121*	*2.00*	*1.038*	*1.006*	*.836*	*131.8*	*109.9*
i	14,173	13,144	1,028	6,807	5,694	2.00	1.038	.997	.834	124.5	104.1
ii	15,064	13,971	1,092	7,321	6,022	2.00	1.038	1.009	.830	132.3	108.8
iii	15,049	13,925	1,122	7,422	6,244	2.00	1.026	1.012	.837	132.5	111.1
iv	15,915	14,673	1,241	7,810	6,543	2.00	1.024	1.005	.842	137.7	115.4
1946	*15,967*	*14,990*	*977*	*7,364*	*6,669*	*2.00*	*1.009*	*.938*	*.844*	*125.9*	*114.0*
i	15,715	14,556	1,157	7,618	6,495	2.00	1.015	.984	.839	132.7	113.1
ii	15,751	14,776	975	7,392	6,580	2.00	1.009	.947	.843	127.2	113.2
iii	16,070	15,168	902	7,285	6,734	2.00	1.007	.913	.844	123.8	114.7
iv	16,331	15,458	873	7,159	6,867	2.00	1.006	.882	.846	120.2	115.3
1947	*16,457*	*15,589*	*868*	*7,043*	*6,989*	*2.00*	*1.001*	*.857*	*.850*	*114.7*	*113.8*
i	16,135	15,293	842	6,962	6,833	2.00	1.000	.863	.847	115.5	113.4
ii	16,021	15,220	801	6,894	6,805	2.00	.997	.858	.847	113.0	111.5
iii	16,565	15,664	901	7,066		2.00	1.001	.854		114.4	
iv	17,107	16,177	930	7,250		2.00	1.004	.851		115.9	

[1] From monthly averages of daily figures, *Banking and Monetary Statistics*, pp. 369–70 and 396, and issues of *Federal Reserve Bulletin* 1942–1948.

[2] For 1929–1947, *Banking and Monetary Statistics*, p. 396, and issues of *Federal Reserve Bulletin*, 1942–1948, with amount for first quarter of 1933 estimated in part from weekly data. For 1919–1928 estimated from (a) required reserves on call dates nearest June 30 and December 31, and (b) ratios of average total deposits other than interbank in reporting member banks for quarters immediately preceding and subsequent to June 30 or December 31 to the amount for the call date, with a slight adjustment on the basis of reported average required reserves for 1929. Data from *Banking and Monetary Statistics*, pp. 132–42, 369–71, 373–75, 387, 395, and 396 and issues of *Federal Reserve Bulletin*, 1919.

[3] Average total reserves adjusted by relatives of percentage requirements, deposit shifting, and deposits not requiring reserves (divided by the first two and multiplied by the third). Annual figures are averages of quarterly figures and may differ from those computed from the average annual relatives. For the difference between columns A and B, 1943–1947, see footnote 6.

[4] Ratio of percentage reserve requirements to such requirements during 1923–1928. Change in percentage requirements during a quarter are treated as in force throughout the quarter, in view of the fact that in most instances the changes were announced in advance. Reduction in reserve requirements for central reserve city banks only in 1942 is not taken into consideration here (see footnote 5).

[5] Proportion of net demand plus time deposits in member banks in each of four categories (net demand deposits in central reserve city banks, net demand deposits in reserve city banks, net demand deposits in country banks, and time deposits), weighted by the relative percentage reserve requirements (13, 10, 7 and 3, respectively, until June 1942, and 10, 10, 7 and 3, respectively, subsequent to that date), converted to relatives with the average for 1923–1928 = 1. These percentages take into consideration the change in reserve requirements in 1942 applicable only to banks in central reserve cities. Deposit figures for 1929–1947 are from monthly or semi-monthly daily averages, and those for 1918–1928 are interpolations from deposits on call report dates (from *Banking and Monetary Statistics*, pp. 73, 81, 87, 93, and 99; annual reports of the Board of Governors of the Federal Reserve System, 1931–1937; and issues of the *Federal Reserve Bulletin*, 1938–1948).

[6] Interpolated from ratio of total deposits adjusted to total net demand plus time deposits in member banks for June 30 and December 31 of each year, converted to relatives with the average for 1923–1928 = 1.0. Data for total deposits adjusted from *Banking and Monetary Statistics*, pp. 34–35, with figures for December 31, 1917–1922, interpolated, and issues of *Federal Reserve Bulletin*, 1942–1948; data for total net demand plus time deposits in member banks from *Banking and Monetary Statistics*, pp. 73–75, and issues of *Federal Reserve Bulletin*, 1942–1948. For 1943–1947, figures in column A refer to data including, those in column B to data excluding, United States government deposits.

[7] Trend values are computed at 1.227 per cent increase per quarter; and are based on the average effective reserves during the 25-quarter period from the 4th of 1922 to the 4th of 1928, inclusive, centered at the middle quarter of that period.

THE ORDERING OF n ITEMS ASSIGNED TO k RANK CATEGORIES BY VOTES OF m INDIVIDUALS

GARRET L. SCHUYLER
Rear Admiral, USN (Retired)

A method is proposed for arranging in order of acceptability a set of n items (research projects, candidates for positions, articles used, viewed, or smelled) on the basis of judgments of m judges. Unlike previous methods, this permits each judge independently to group in any manner those items tied for his favor, and to omit voting on any items concerning which he cannot decide. The sampling probabilities for estimating accuracy are not here considered.

INTRODUCTION

WE MAY STATE in the following most general terms, the problem faced by a panel of judges ranking items, projects or persons according to some immeasurable quality.

PROBLEM

Items $P_1, P_2, \cdots, P_n$ are to be arranged in order by votes of individuals $M_1, M_2, \cdots, M_m$. Let u_i be number of items concerning which individual M_i is too undecided to express his opinion, and $q_i, r_i, s_i, \cdots$ numbers of items he assigns to categories $a, b, c, \cdots$ of decreasing worthiness. Then upon what criterion C_j is the position of item P_j among the other items best based?

SOLUTION[1]

Consider a table in which each column corresponds to an item and each row to a voter, with letters such as $a, b, c, \cdots$ to indicate the ratings of the several projects, or u for "undecided." The letters $a, b, c, \cdots$ are to be replaced by numbers in the same relative order, and the totals of the columns are to be taken to indicate the order of excellence to be assigned as a composite judgment to the items.

---

[1] M. G. Kendall: The Advanced Theory of Statistics. Griffin, 1943, p. 410.

For this purpose the first rank numbers to suggest themselves are the midranks

$$(1/2)(q_i + 1),\ q_i + (1/2)(r_i + 1),\ q_i + r_i + (1/2)(s_i + 1),\ \cdots$$

in the blocks of items receiving a-votes, b-votes, c-votes, . . . respectively of member M_i. But the average of the $(n-u_i)$ rank numbers which M_i assigns on this basis is only $(1/2)(n+1-u_i)$; and this function of the number of u-votes varies from member to member.

To have rank numbers from all members' votes properly additive, the average rank numbers for all members' votes should be equal. To effect this it seems simplest to multiply each member's midrank numbers by an adjustment factor

$$A_i = \frac{n+1}{n+1-u_i} \tag{1}$$

which will make the average rank numbers for a-votes, b-votes, etc. of each member equal to $(1/2)(n+1)$. Also, a rank number $(1/2)(n+1)$ should be assigned for each u-vote.

If R_{ai}, R_{bi}, R_{ci}, etc. and R_{ui} are the rank numbers thus assigned to items receiving respectively a-votes, b-votes, c-votes, etc. and u-votes of member M_i, we thus have

$$\begin{aligned}
R_{ai} &= A_i(1/2)(q_i + 1),\\
R_{bi} &= A_i(q_i + (1/2)(r_i + 1)),\\
R_{ci} &= A_i(q_i + r_i + (1/2)(s_i + 1)),\\
&\cdots\cdots\cdots\cdots\\
R_{ui} &= (1/2)(n + 1).
\end{aligned}$$

The criterion for item P_j is then the sum C_j of the rank numbers R of the votes of all members on this item.

FLEXIBILITY OF THE METHOD

These equations do not restrict panel members to the use of any particular number of categories or ranks, nor do they even require that all members use the same series of ranks in the same voting. A member may or may not leave the ratings of individual items undecided.

The minimum possible number k of different ranks which a member may meaningfully assign is obviously 2. When all members use the

maximum value n of k, the order of the C_j is the same as that indicated by Kendall[2] without consideration of the more general problem here described.

THE CASE $k=3$

Principal interest seems most likely to lie, however, in less idealized cases where, unlike judges encountered in the textbook examples, panel members will not

(a) agree to vote on every item (i.e., to use no u-votes), nor
(b) agree to avoid tie votes (i.e., to use as many ranks as there are items).

Recently, for example, panels of 15–20 voters were required to arrange series of 40–122 problems in order of worthiness for retention on a possibly reduced research budget. Few members had specialized on all parts of the field covered. Nearly every member encountered items on which he felt himself insufficiently well qualified to cast a possibly decisive vote.

It was soon agreed that members might withhold opinions on any items with which they felt themselves too unfamiliar; and their comments showed certain members of some panels would cast only u-votes on a majority of the items.

Again very unlike those most obliging judges encountered in textbook examples, these panel members declined to use dozens of ranks. The most they could be persuaded to use was only three ranks (corresponding perhaps to rough classifications of "very worthy," "worthy" "less worthy").

As a practical illustration consider Table I in the first four blocks of which are shown (a) identifying numbers of voters, (b) their votes on two very closely competing items P_{33} and P_{28}, (c) their numbers of votes on all items, and (d) rank numbers, and their averages, for items P_{33} and P_{28}.

CRITERIA WHEN $k=3$

In spite of its comparative neglect in textbooks this case in which no voter discriminates more than three ranks is sufficiently common to have developed an approximate or "1, 2, 3 method" for its treatment.

---

[2] For completeness one must mention also the possible, though I think less promising methods, by which the order of items (a) is so selected as to minimize the number of "inversions" of rank order in relation to ranks assigned by the various members' votes or (b) is taken in accordance with *net* numbers of votes which rank each item superior to other items or (c) is based upon "Scores for Ranked Data," Table XX, Fisher and Yates' "Statistical Tables for Biological, Agricultural, and Medical Research" (1943).

TABLE I

ACTUAL DATA OF VOTES ON ITEMS P_{32} AND P_{33} OF A 41-ITEM SERIES VOTED ON BY A 15-MEMBER PANEL

M	Votes on Items P_{32}	Votes on Items P_{33}	All Votes a	All Votes b	All Votes c	All Votes u	Rank No's. By C_j/n P_{32}	Rank No's. By C_j/n P_{33}	"Scores" P_{32}	"Scores" P_{33}	Rank No's. By C_j'/n P_{32}	Rank No's. By C_j'/n P_{33}
1	c	c	16	9	16		33.5	33.5	3	3	34.7	34.7
2	c	b	11	16	13	1	35	20	3	2	34.7	21
3	b	b	16	16	9		24.5	24.5	2	2	21	21
4	b	b	9	22	10		20.5	20.5	2	2	21	21
5	c	u	9	13	9	10	35.4	21	3	—	34.7	—
6	b	b	22	11	8		28	28	2	2	21	21
7	b	c	22	13	6		29	38.5	2	3	21	34.7
8	b	c	11	17	13		20	35	2	3	21	34.7
9	c	b	15	15	11		36	23	3	2	34.7	21
10	c	c	13	13	15		34	34	3	3	34.7	34.7
11	a	b	17	7	17		9	21	1	2	7.3	21
12	c	b	15	16	10		36.5	23.5	3	2	34.7	21
13	c	c	15	12	14		34.5	34.5	3	3	34.7	34.7
14	b	b	18	14	9		25.5	25.5	2	2	21	21
15	b	c	24	9	8		29	37.5	2	3	21	34.7
			Means				Means		Means			
			15.5	13.5	11.2	.7	28.7	28.0	2.40	2.43	26.4	26.7
		Max	24	22	17	10	±5.9	±5.9	±.63	±.49	±7.6	±6.7
		Min	9	7	6	0	(m.d.)		(m.d.)		(m.d.)	

Votes	a	b	c	u
Item P_{32}	1	7	7	0
Item P_{33}	0	8	6	1

Example: For judge M_5

$$A_5 = \frac{41+1}{41+1-10} = 1.31$$

$$R_{c5} = 1.31(9+13+\tfrac{1}{2}(9+1)) = 35.4$$

$$R_{u5} = \tfrac{1}{2}(41+1) = 21$$

With this the a-votes, b-votes and c-votes are respectively assigned the arbitrary weights 1, 2, 3 and items are ranked according to weighted averages.

When a_j, b_j, c_j are numbers of a-votes, b-votes, and c-votes respectively which item P_j receives from all members, the criterion often used is

$$C_j' = \frac{a_j + 2b_j + 3c_j}{a_j + b_j + c_j}.$$

This gives an approximation to the order specified by the C_j in that, if each voter places exactly one-third of the items in each category, the order among the numbers C_j' is the same as that among the C_j. In other cases the order may differ.

The ordering of the items according to C_j' has the dubious property of making a c-vote and an a-vote count always the same as two b-votes for an item. Furthermore, this method makes a u-vote count the same as the average of all other votes on the *same* item, whereas ordering by the C_j makes a u-vote count the same as the average of the votes on *all* items.

With moderately large numbers of voters these differences cancel out surprisingly well. For example, in the voting on 41 items by a 15-member panel cited above, there was only one pair of items whose relative rank differed by the two methods. These were the items P_{28} and P_{33} for which the detailed vote is shown in Table I. Even in this single case the margin can doubtless be ascribed to chance.

CONCLUSIONS

The simple and rough ordering according to C_j', carried out to the nearest tenth of a unit, and with differences of less than two-tenths of a unit between pairs of items disregarded, will probably meet most accuracy requirements encountered in practice, at least if the proportion of undecided votes is as low as in Table I. But the better method using C_j involves only a negligible amount of additional labor and, if only because of its extreme generality and flexibility, it is believed that it should be more widely understood.

The degree of consistency in data of this sort may be examined in various ways, and probability calculations are essential in examining it. Interest in this aspect of the problem, and the appropriate methods, may vary widely according as the panel is intended (a) to furnish a sample of the opinion in some very large and not too select, population, or (b) to furnish a record only of its own collective judgments.

LEVELS OF SIGNIFICANCE FOR VARIANCE RATIO OF TWO SAMPLES OF EQUAL SIZE

C. J. KIRCHEN*
University of Michigan Aeronautical Research Center, Ypsilanti, Michigan

THIS CHART permits two-tailed probability evaluations of ratios of variances of two equal-sized samples at other than the conventional points of 5 and 1 per cent which are 10 and 2 per cent two-tailed levels, respectively. The chart was plotted from the tables of Thompson and Merrington (*Biometrika:* 33, pp. 73–88), the 20 per cent table of H. W. Norton (*Statistical Tables for Biological, Agricultural and Medical Research*, by Fisher and Yates; Oliver and Boyd, Ltd., Edinburgh and London), the .1 per cent table of Colcord and Deming (*Sankhya*, ii, pp. 423–424). The plotted points for 19 and 49 degrees of freedom were obtained by harmonic interpolation on n_1, linear interpolation on n_2.

The reader who is interested in acquainting himself with the mathematical basis for this chart is referred to *Introduction to Mathematical Statistics* by P. G. Hoel; John Wiley and Sons, Inc., New York; 1947; pp. 152–154.

* Formerly with Remington Arms Company, Inc., Bridgeport, Conn.

Corrections to article published in Volume 43, No. 242, June 1948

ON ESTIMATING PRECISION OF MEASURING INSTRUMENTS AND PRODUCTS VARIABILITY

FRANK E. GRUBBS

Page 244, middle of page: Change $x_1,\ x_2, \cdots x_1 \cdots x_n$ to read $x_1,\ x_2, \cdots x_i, \cdots x_n$

Page 245, second line from bottom: Change $x_i + \bar{x}$ in second term to read $x_i - \bar{x}$

Page 245, bottom line: change c_1 to c_i.

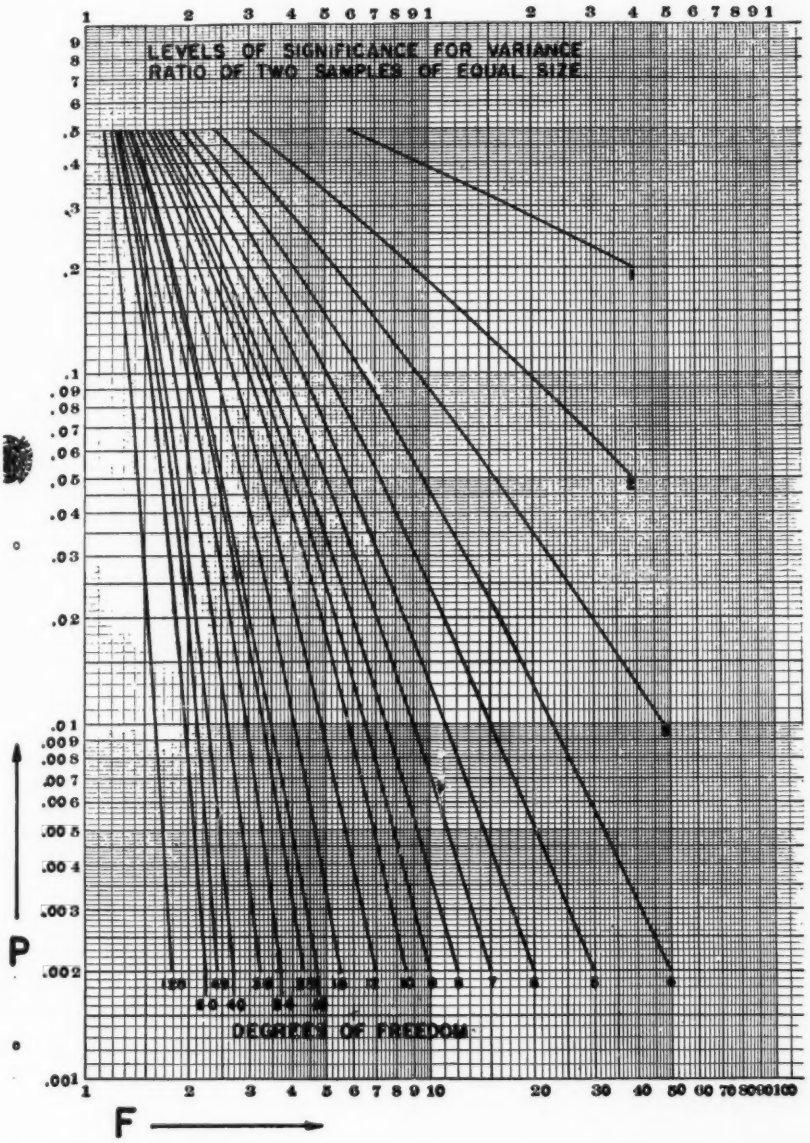
Page 249, line 8: Change x_1 to x_i and change x_{i1} to x_i

Page 253, tenth line from bottom of page: Change $e_{11} + e_{12}$ to $e_{11} - e_{12}$

Page 261, formula (30): Change e_n to e_N

Page 264, line 17: Change the word "variance" to read "variation"

Page 264, footnote, third line from bottom: Change "use" to read "case"



The author wishes to acknowledge with thanks the permission granted by the Remington Arms Co. to publish the attached chart, which was devised in connection with work in which the author was engaged with the company.

MAIN EFFECTS AND INTERACTIONS

D. J. FINNEY
University of Oxford

When the effects of two factors are simply additive, choice of estimates can be made entirely on grounds of minimal variance. When interaction exists, estimation of the magnitudes of main effects and interactions requires an arbitrary definition of what constitutes a main effect. Whether the class frequencies be equal, proportionate, or disproportionate, the definition must conform to the use that is to be made of the results and must not be dependent merely on the convenience of an analysis of variance.

THE analysis of variance is now so widely recognized as one of the most powerful techniques for the examination of statistical data that its users are sometimes in danger of regarding it as an end in itself instead of as a means to an end. In a well designed experiment, for example, the analysis of variance of the results and the summarizing of the effects of treatments are closely related and may even be merely different aspects of the same set of calculations; nevertheless, few things betray the inexperienced statistician more readily than a triumphant presentation of an elaborate analysis of variance table coupled with an almost complete neglect of any statement of treatment means. With fortuitous or observational data, in which symmetry may be less marked than in data from a planned experiment, analysis of variance is still a valuable method of studying different components of variation, but additional care is needed in order to ensure that the pattern of analysis adopted is appropriate to the data and to proceed satisfactorily from analysis to summary of results.

In a recent letter to *Nature*, Vajda [1] has drawn attention to a difficulty that occurs in the definition of a "main effect" for an analysis of variance of data with unequal sub-class frequencies. Proper understanding of his point is important even with data having equal frequencies in the sub-classes, for only when the arbitrary nature of any definition of a main effect, at least if interactions are present, is appreciated can the analysis of variance and the summarizing of conclusions be seen in right perspective. Yates [2] has corrected certain misconceptions that might arise from an incautious reading of Vajda's letter. A fuller statement, however, may be valuable to some whose knowledge of the analysis of variance is based more on text books of statistics than on their own experience.

The problem may be illustrated by reference to data on the interaction of two factors. Suppose that n_{ij} observations are available on a variate x, at a level i of factor A and a level j of factor B in combination, and that the mean yield (using "yield" as a general term for the variate under analysis) for this combination A_iB_j is x_{ij}. Suppose further that p levels of A and q levels of B have been examined; if certain combinations of levels are lacking, some of the n_{ij} will be zero. The mean yields may be set out as a $p \times q$ table:

		Level of A			
	j \ i	1	2	3 $\cdots$	p
	1	x_{11}	x_{21}	x_{31} $\cdots$	x_{p1}
Level of B	2	x_{12}	x_{22}	x_{32} $\cdots$	x_{p2}
	3	x_{13}	x_{23}	x_{33} $\cdots$	x_{p3}
	.	.	.	.	.
	.	.	.	.	.
	.	.	.	.	.
	q	x_{1q}	x_{2q}	x_{3q} $\cdots$	x_{pq}

and the sum of squares of deviations from the general mean may be partitioned as follows:

	d.f.
Between classes	$pq - 1$
Within classes	$S_i S_j n_{ij} - pq$
Total	$S_i S_j n_{ij} - 1$

where S_i, S_j indicate summations over all values of the suffixes i, j, respectively.[1] This much is easy: difficulties arise only when attempts are made to examine particular components of the sum of squares between classes, and especially those components usually described as "main effects."

The effect of factor A is said to be independent of factor B if the difference in mean yield between any two levels of A is the same at each level of B, except for variations attributable to random sampling: $(x_{ij} - x_{i'j})$ is dependent upon i, i' but, apart from sampling variation, independent of j. That differences in mean yields between any two

[1] The "total" and "between classes" degrees of freedom will be reduced by 1 for every zero value of n_{ij}.

levels of B are similarly independent of the level of A necessarily follows. The *main effect* of A is then naturally defined as the set of differences between levels of A, and will be estimated by fitting a set of $(p-1)$ constants for A simultaneously with a set of $(q-1)$ constants for B, using the method of least squares (Yates [4]) and weighting each x_{ij} with n_{ij}.

In a planned experiment, n_{ij} will often be a constant, and the subdivision of the sum of squares between classes into main effects and interaction is then very simple. (Further components for "blocks" or other constraints may have to be removed from what has been described above as "within classes," before a residual error mean square can be formed, but that does not affect the principles discussed here.) The required main effects are the differences between simple arithmetic means of columns and rows in the x_{ij} table. When n_{ij} is not constant, in general the analysis is more complicated, but the particular case of *proportional frequencies*, mentioned by Vajda, can be treated simply. If

$$n_{ij} = u_i v_j,$$

where u_i is independent of j and v_j is independent of i, differences between column-means weighted according to frequencies, namely

$$x_{i.} = \sum_j n_{ij}x_{ij} / \sum_j n_{ij} = \sum_j v_j x_{ij} / \sum_j v_j$$

are easily seen to be orthogonal with the similar set of contrasts based on row means

$$x_{.j} = \sum_i n_{ij}x_{ij} / \sum_i n_{ij} = \sum_i u_i x_{ij} / \sum_i u_i.$$

Furthermore, in virtue of the well-known property of a mean of observations weighted inversely as their variance, differences between the $x_{i.}$ estimate the main effect of A with lesser variance than would any other method of averaging the columns. Similarly the $x_{.j}$ give the most efficient estimate of the main effect of B. It is therefore natural to regard the $x_{i.}$ and the $x_{.j}$ as providing the best estimates of the main effects from the data. The least squares procedure of fitting constants leads to exactly the same results. An analysis of variance, using the familiar formulae to give sums of squares for main effects, leaves a third constituent of the total sum of squares between classes, with $(p-1)(q-1)$ degrees of freedom, which is appropriate to a test of significance of the null hypothesis of independent action of A and B;

this is usually described as a significance test on the interaction, $A.B$.

The difficulty to which Vajda draws attention arises only when the null hypothesis of independence has to be abandoned. The effects of the two factors are then no longer simply additive: the definition of the main effect of A, and in particular of the difference between mean yields for levels i and i', must therefore be re-examined. This difference now varies in some manner dependent upon the level of B, and, even though $n_{ij} = u_i v_j$, considerations other than minimal variance may have to determine the estimated average value for the data. To use $x_{i.}$ as before is tempting, on account of its computational convenience and analytical elegance, for it still leads to the weighted average of quantities $(x_{ij} - x_{i'j})$ having minimal variance and it still gives a neat analysis of variance table. This is the estimation termed by Vajda that of "proportionate frequencies"; he appears to prefer to use mean values for levels of A and B defined by

$$\bar{x}_{i.} = \underset{j}{S}\, x_{ij}/q, \qquad \bar{x}_{.j} = \underset{i}{S}\, x_{ij}/p,$$

the simple averages of class means, identical with the $x_{i.}$ and $x_{.j}$ only if n_{ij} is the same for all i, j. This he speaks of as the "method of fitting constants," but he uses the phrase in a sense rather different from Yates, who reserved it for the estimation of additive main effects when interactions were ignored. Vajda's usage implies the "estimation" by least squares of $(pq-1)$ parameters from the $(pq-1)$ differences between class means x_{ij}, so giving a perfect fit.

Yates [5] points out that "Inasmuch as the frequencies actually obtained are likely to be similar to those generally occurring in the type of material under consideration, the definition of the main effects implied by the use of proportionate frequencies is usually more appropriate in practice than that based on equal frequencies." Whether in any one instance the definition is "more appropriate" depends upon the particular circumstances, and elegance of analysis alone must not be the criterion. Unless the v_j are proportional to frequencies for the q levels of B specially interesting to the experimenter (as when they are relative frequencies in a natural population of which the experimental material is intended to be representative), there is no good reason for preferring $x_{i.}$ to $\bar{x}_{i.}$. The point to be emphasized is that both are arbitrary: $\bar{x}_{i.}$ arbitrarily assigns equal importance to each of the q values of $(x_{ij} - x_{i'j})$, whereas $x_{i.}$ arbitrarily weights these differences in proportion to v_j. Conceivably, though rarely, the mean values of A required in summarizing the data would be defined by some extraneous

system of weighting for levels of B, using weights v_j^* independent of i, but different from v_j, to give

$$x_{i.}^* = \underset{j}{S} v_j^* x_{ij} / \underset{j}{S} v_j^*.$$

These, and a similar set of $x_{.j}^*$ for B, would estimate two sets of main effects mutually orthogonal and themselves defining the interaction as the residual set of $(p-1)(q-1)$ contrasts. As a method of summarizing the results, this would be quite legitimate, though it would not correspond to a simple analysis of variance. Vajda rightly stresses that "the definition of an effect or of an interaction should not depend upon whether the frequencies are noticed to be such that a change in the definition leads to easier computations," but mistakenly concludes that "the method of fitting constants should therefore always be accepted as the only legitimate one." When the interaction is not negligible, estimation of the main effect of A must depend upon circumstances, and no definition, whether $x_{i.}$, $\bar{x}_{i.}$, or the more general $x_{i.}^*$, can be claimed as "the only legitimate one." The problem forces itself upon the attention for the case of proportionate sub-class frequencies because alternative procedures are then very evident. It still exists at the extremes of symmetry and asymmetry. The wise investigator will take account of this, when possible, by choosing his sub-class frequencies in accordance with the proportionate frequencies condition, so as to combine maximal precision of estimation with the weighting system appropriate to his interests.

When all the n_{ij} are equal, as in the majority of planned experiments, the assumption that the simple average for each level of A over all levels of B gives the ideal estimate of the main effect of A is sometimes made too readily; if there is an interaction between A and B, the interpretation of these averages needs to be most carefully considered. A common situation is that in which the "levels" of factor B represent different methods of performing some operation. If the effect of A depends upon which method is chosen, no type of average A effect can be of much interest unless the q methods tested in the experiment either are the only possibilities or are in some way representative of all possible methods. An A effect averaged over an entirely arbitrary selection of levels of B may give a formal completeness to a summary table of results, but is then meaningless or even misleading: the results should instead be interpreted in terms of the effect of A for each B_j separately. For example, in a field experiment on a root crop, factor A might represent amounts of nitrogenous fertilizer and factor B different forms of

cultivation (inevitably a small selection only from all the methods of cultivation that might interest the experimenter). If analysis shows the hypothesis of simple additive effects to be untenable, a nitrogen "main effect" averaged over several cultivation treatments may be little use except as a broad qualitative expression of the results. Any detailed interpretation of the experiment will need to examine either responses to nitrogen under each system of cultivation separately or differential cultivation effects at each level of the fertilizer separately.

When $n_{ij} \neq u_i v_j$, no simple analysis of variance exists to tempt the experimenter to sacrifice meaning for elegance, but again he must not assume that the unweighted averages $\bar{x}_{i\cdot}$ and $\bar{x}_{\cdot j}$ are necessarily what he requires. In the absence of interaction, the estimation problem is always simple in principle, and choice between alternatives can be based on their relative precisions. In the presence of interaction, the definitions of main effects and interactions are inextricably bound together; exact meaning can be given to interaction only as the departure from a specified system of main effects, and definition of the estimates of the latter must involve, explicitly or implicitly, arbitrary weighting of the effects of one factor at different levels of another.

The above discussion has been restricted to two factors and their interaction, but the generalization to a greater number should be immediately apparent. This note is concerned only with the logical basis for speaking of main effects and interactions; the technique of estimating them from non-orthogonal data, and of fitting them into the framework of analysis of variance, has been described by Yates [3, 4] and, in great detail and with wealth of illustrative examples, by Snedecor [1].

REFERENCES

[1] Snedecor, G. W., *Statistical Methods* (4th edition), Iowa State College Press, Ames, Iowa, 1946.

[2] Vajda, S., "Technique of the Analysis of Variance," *Nature*, London, Vol. 160, 27, 1947.

[3] Yates, F., "The Principles of Orthogonality and Confounding in Replicated Experiments," *Journal of Agricultural Science*, Vol. 23, 108–145, 1933.

[4] Yates, F., "The Analysis of Multiple Classifications with Unequal Numbers in the Different Classes," *Journal of the American Statistical Association*, Vol. 29, 51–66, 1934.

[5] Yates, F., "Technique of the Analysis of Variance," *Nature*, London, Vol. 160, 472–473, 1947.

A TEST FOR SYMMETRY IN CONTINGENCY TABLES

ALBERT H. BOWKER
Stanford University

SUPPOSE THAT a sample of N individuals is classified according to two arguments, A and B, in an m by m contingency table and suppose that the categories of classification A are associated in some unique way with the categories of classification B. For example, in the study in which this problem arose the categories were related to the position of plates of stickleback (Gasterosteus) and the two arguments corresponded to right and left sides; it was logical to associate identical patterns on opposite sides. Denote by P_{ij} the probability that a randomly chosen individual belongs to the ith row and jth column of the table. As usual, a dot indicates summation over the replaced index.

DISTRIBUTION OF POSSIBLE EVENTS

	Category	Classification A						Marginal Distribution
		1	2	—	—	—	m	
Classification B	1	P_{11}	P_{12}	—	—	—	P_{1m}	$P_{1.}$
	2	P_{21}	P_{22}				P_{2m}	$P_{2.}$
	—	—	—				—	
	—	—	—				—	
	—	—	—				—	
	—	—	—				—	
	m	P_{m1}	P_{m2}				P_{mm}	$P_{m.}$
Marginal Distribution		$P_{.1}$	$P_{.2}$				$P_{.m}$	

It is desired to test the hypothesis that the above table is symmetric (i.e., $P_{ij}=P_{ji}$). If this hypothesis is true, then $P_{k.}=P_{.k}$; that is, the marginal distribution of individuals is the same for classifications A and B. Further, the conditional distribution with respect to A for the kth category of B is equal to the conditional distribution with respect to B for the kth category of A. This equality of conditional distributions in general implies a stronger equivalence between the classifications than equality of marginal distributions, although for $m=2$ equality of marginal distributions implies complete symmetry. The weaker hypothesis of equality of marginal distributions would also be of interest, especially in the absence of symmetry; this problem appears to be somewhat more difficult to handle by straightforward methods.

The proposed test for the hypothesis of symmetry is of the familiar χ^2 type. Let n_{ij} be the observed frequency in the ith row and jth column of a table of observed quantities arranged as above. Then

$$(1) \qquad \sum_{i>j} \sum \frac{(n_{ij} - n_{ji})^2}{n_{ij} + n_{ji}}$$

has for large N the Chi-square distribution with $m(m-1)/2$ degrees of freedom. A large value of χ^2 leads to rejection of the hypothesis. Note that the value of χ^2 is computed from non-diagonal frequencies, which is reasonable since diagonal frequencies should not afford evidence tending to disprove the hypothesis of symmetry.

In cases in which the observed frequencies are small and no regrouping of frequencies is possible (e.g., in a 2×2 table) an exact test may be of interest. Specifically, in a 2×2 table an exact test may be made by testing with any table of the binomial distribution the hypothesis that n_{12} and n_{21} are observed "successes" and "failures" in a random sample of $n_{12}+n_{21}$ events. From equation (1) we obtain for a 2×2 table the following χ^2 statistic with one degree of freedom:

$$(2) \qquad \frac{(n_{12} - n_{21})^2}{n_{12} + n_{21}} .$$

This is equivalent to the value obtained by a test of goodness of fit of a binomial with $P=\frac{1}{2}$. Formula (2) was obtained by McNemar [1], who lists applications in psychological testing. A problem in the same field prompted a recent query in *The American Statistician* [2].

Several intuitive derivations of the above test may be given, but a rigorous justification may be obtained from a theorem of Cramér [3], page 426, which puts the use of χ^2 tests in contingency tables on a sound theoretical basis. This procedure affords an interesting example of a routine application of the general theory and might well be an exercise for a course in the theory of χ^2 tests. After verifying certain regularity conditions, the method of Cramér can be applied by computing the quantity

$$(3) \qquad \sum_{i} \sum_{j} \frac{(n_{ij} - NP_{ij})^2}{NP_{ij}}$$

where the P_{ij} are replaced by their maximum likelihood estimates under the hypothesis of symmetry. These estimates turn out to be $n_{ij}+n_{ji}/2N$ which yield equation (1) on substitution into (3).

An alternate derivation is available if (1) is regarded as a sum of $m\ (m-1)/2$ terms of type (2). This approach requires the assumption that we are sampling from a population with fixed sums for all symmetric pairs of cells. While both methods lead to the same result, the more general one based on Cramér's theorem seems slightly preferable. Until the test procedure is derived, it is not obvious that the use of the conditional distribution with fixed totals for symmetric pairs entails no loss of efficiency.

The problem which prompted this note was suggested by Professor Frank Weymouth.

REFERENCES

[1] McNemar, Quinn, "Note on the sampling error of the difference between correlated proportions or percentages," *Psychometrika*, Vol. 12, No. 2, 1947, pp. 153–157.

[2] "Question 1," *American Statistician*, Vol. 1, No. 2, 1947, p. 12.

[3] Cramér, Harald, *Mathematical Methods of Statistics*, Princeton University press, 1946.

THE WAR PRODUCTION BOARD'S STATISTICAL REPORTING EXPERIENCE

PART IV

QUESTIONNAIRE UNITS OF MEASURE AND NOMENCLATURE

DAVID NOVICK
University of Puerto Rico
AND
GEORGE A. STEINER
University of Illinois

Effective wartime administration of the industrial system required the use of a common factual language with which government and industry could communicate. No such system existed in the prewar period. An effort to develop a suitable quantitive language was begun early in the war, and met with some of the most difficult problems encountered by wartime management. Although substantial progress was made, a stabilized and commonly understood system of communication was never attained. This article discusses some of the outstanding problems in connection with questionnaire units of measure and terminology. An analysis of the development of material and product classification and terminology is followed by a discussion of the methods employed to standardize concepts relating to general terminology in questionnaire instructions, to units of measure, to reporting time periods and to end-product usage.

PRODUCT NOMENCLATURE AND CLASSIFICATION

THE PROBLEM of developing uniform identification and classification of materials and products was more than a question of simple definition. Standard definitions of both products and materials were available. So also were classification systems. The problem lay in widely varying practices of industry relating to record-keeping methods and definitions, procurement patterns, and commonly used day-to-day expressions. Into this traditional lack of uniformity were injected rapidly changing wartime influences, new products and materials, and substantial modifications and revisions of old products. These conditions lay at the base of the broad problem of unifying product and material concepts.

Product names and other words used to identify them early in the war meant various things to different minds and by no means did they

always carry either the dictionary meaning or the technical one given to them in industry. Under the operation of the early preference rating system, products and materials for which preferred treatment was extended were defined in broad terms, such as bombs, cranes, explosives, and tin. It was found very soon that manufacturers in reporting like products under this system used so many different names and so many varying classifications that identification for tabulation, allocation, and issuance of preference ratings became exceedingly difficult.

As supply relative to demand tightened for an increasingly large number of products and materials, and as administrators became more familiar with their products and the problems of controlling them, a drive began for greater precision of definition. But getting more detail intensified the problem of nomenclature since the result was often a further entanglement in varying industry usage and variations of a given product in grade, size, weight, cost and other characteristics, such as horsepower in the case of certain machines, warmth-conserving qualities in the case of waterfowl feathers, or wear-resistance in the case of leather and textiles.

Relatively little improvement was noted in more precisely defining products and materials given in the *Priorities Critical List* to which preferred treatment could be extended on rating certificates. The move for more precision really began with the introduction of blanket ratings authorized by P Orders in the spring of 1941. The earliest P Orders were couched in vague terms with respect both to the product to which preference ratings might be applied and to the materials and components to which ratings might be extended. Continuous and often voluminous successive amendments, however, brought greater and greater detail into terminology.

Order P-2 illustrates this evolution. Dated March 26, 1941, this order related vaguely to "material for the production of machine tools." No further definition of machine tools was given. But the list of 11 materials to which rating could be extended to facilitate machine tool production with "Machine Tools, consisting of power-driven metal working machines, hydraulic and mechanical presses, and welders," and continued with components such as "motors and other electric equipment," "abrasives," "machine parts and equipment," etc. In the next month a series of amendments more precisely defined equipment, viz., "power-driven metal-working machines and testing equipment," "drill presses," "tornio and automatic screw machines," "nibblers and cutoff machinery," "magnetic chucks and grinding machines," and so on. With each successive amendment the list of specific tools became longer and more precise.

Reporting forms naturally followed the trends established by the orders to which they were attached. Early reporting forms provided blank spaces for manufacturers to report products coincident with the move for more precise identification. Later forms often preprinted on the form itself the exact products to be reported. This development, obviously, served not only to limit reporting to products specified but also facilitated uniformity of response and understanding.

Limitation orders, first issued in mid-1941, generally described the products to which they applied in far more precise terms than early P Orders. This was due partly to experience in connection with P Orders and partly because the L Orders, as an instrument for limiting production of a specified item, had to be precise to insure compliance. In addition, of course, the limiting of production necessitated considerable care because it moved deeply into the entire problem of industrial profits and conversion to war work.

The definition of trucks and motor vehicles set forth in L-1, issued in August 1941, did not change materially in successive amendments to the order or to other L Orders issued to apply to model types originally incorporated in L-1. L-26, which was the next to last L Order issued in 1941, restricted the production of farm machinery, farm equipment, and parts. Products covered by the order were listed in great detail. The reporting instrument first used under the order, PD-284, preprinted each of the approximately 175 specific pieces of machinery and equipment on the stub of the form. This sort of pattern was followed by many L Orders and related forms, such as L-7 for refrigerators and its forms PD-176, 177, and 417; L-17 for vacuum cleaners and its forms PD-170, 171 and 417; and L-21 for automatic phonographs and its related forms PD-182 and PD-417.

Experience of administration brought greater improvement in product definition. But among the many orders, regulations and questionnaires, there existed great variation in definitions, a situation which although constantly improved lasted throughout the war period. Attempts of the WPB Requirements Committee and the various WPB Divisional Requirements Committees to direct product production and distribution and to allocate individual products under control were constantly impeded by these conditions. Adding and duplicating machines, for instance, as used in Service Equipment Division in early reports, included both electric and non-electric types. Sewing machines, on the other hand, were reported under the headings: "electric," "hand," and "treadle." Benzene as originally reported by the Chemicals Bureau meant both the pure pharmaceutical grade and the much lower motor grade. Hot water storage tanks defined by the Plumbing

and Heating Division included four products ranging in weight of contained carbon steel from 50 to 1700 pounds, while track laying tractors, on the other hand, were subdivided into five weight classes and two models. Domestic electric mechanical refrigerators were reported by the Consumers Durable Goods Division in terms of units, regardless of size or weight. Such lack of system was widespread and was justified by administrators as an essential compromise of administration.

The need for better definitions of individual products coincided with the demand for establishing a usable classification to cover the entire spectrum of individual manufactured products. The problem of obtaining better definitions and classifications of individual products and the industries which manufactured them was not new to government statisticians. Government agencies such as the Bureau of the Census, the Bureau of Labor Statistics, and the Bureau of Foreign and Domestic Commerce had devoted considerable attention to the problems of properly classifying industrial data. Each governmental agency, however, maintained its own classification system and a given product or establishment was often classified so differently that data collected by one could not be related to that of the other.

To grapple with this problem in the prewar years, a committee was formed in 1934 under the sponsorship of the Central Statistical Board. The work of this group resulted in the publication in 1937 of the "Standard Industrial Classification."[1] When this volume was completed a "Standard Materials Classification" was begun but was not available in the early defense period when emergency requirements for standard materials classifications became important. By 1941, classification was a problem of considerable magnitude in the work of the OPM. In its efforts to create meaningful summaries of supply contract awards and the volume of preference rating certificates the "Standard Industrial Classification," as modified in the 1939 Biennial Census of Manufacturers, was the classification system used.

The groupings used in the Biennial Census of Manufacturers, however, did not provide for an adequate coverage of all the poducts involved in wartime production. As a result, the Division of Statistical Standards of the Bureau of the Budget (successor to the Central Statistical Board), with the aid of the WPB, The Bureau of the Census and other federal agencies represented on the Federal Committee on Economic Statistics, undertook a cooperative effort which resulted in the development of the "Defense Program Product Classification."

[1] At the end of the article.

This system subdivided many of the broad categories previously used and added new classifications for war products and their complementary items. Although the "Defense Program Product Classification" provided a foundation for the summarization of economic data, it was deemed necessary to group the data further to convey information to top executives. As a result, when the first PD-275 was issued in January 1942, the established classifications were grouped into 111 new product groups covering manufacture in the metal-working industries. In addition, seven new groups were established covering such activities as shipyards, naval air stations, and proving grounds.

Following discontinuance of form PD-275 this classification system was expanded and modified to tabulate form PD-25A, used under the first comprehensive metal allocation system adopted by the WPB. Under the product classification system used for PD-25A it was possible to classify production data with the specific product to which they related, to summarize data to reflect wartime industrial operations, and to provide summary information which the WPB could use for product administration.

Despite the great improvement in product classification for tabulation purposes, considerable difficulty arose in the WPB in attempting to use PD-25A data for administrative purposes. The core of the problem rested in the fact that the WPB internal assignment of product responsibility varied considerably from the product arrangements used in tabulation. In addition, WPB administrators felt that they needed more detail than it was possible to get from manufacturers' PD-25A schedules. Both problems could have been solved, but for these and other reasons this classification was seldom used for day-to-day administrative purposes following the discontinuance of the PRP and its form PD-25A in the spring of 1943. For tabulating such comprehensive general policy forms as WPB-732, CMP-7 and WPB-517, however, this system, continuously modified and improved, remained in use throughout the war period.

With the introduction of the Controlled Materials Plan in the spring of 1943 a new product classification had to be devised to fit its operational needs. Products were divided into two groups: Class A and Class B. Class B products originally were those in the component class but eventually included end-products as well. These were products the controlled material allotments for which were made by the WPB. All other products were Class A products and for the most part allotments were made to them through Claimant Agencies. Over and above the

Class A products were Claimant Agency programs, such as tanks, destroyers, and submarines. Both the Class A and Class B specifications of products were designed essentially for administrative purposes without regard to the logic of product nomenclature.

Material Nomenclature and Classification

The evolution of material definitions and classification paralleled that for products. Efforts to develop adequate information about basic materials brought to a head confusion which had existed as far back as the early NDAC days. There was little consistency of ideas about what constituted a basic material or what shape and form of elements composed it. Each industry had its own definitions. So did early war administrators.

Even a basic concept such as steel ingot was frequently misunderstood, misapplied and misused. Economists and statisticians, for example, in talking about steel thought fundamentally in terms of ingot and spoke continually of the capacity of American steel producers as being 60 million tons of ingot per year. In contrast, the steel industry men talked about finished steel, for example, plate, structural shapes, bars, sheets, etc., but knew that in their conversation tons of ingot meant 75 per cent or less of comparable tons of finished steel shapes and forms. Various operating divisions of the early war agencies were staffed with steel consumers who visualized tons of steel as finished products. They spoke of steel in material tons which translated ingot into finished fabricated component or end products. In this sense an ingot, depending upon the product, meant much less than 75 per cent of comparable basic ingot tons. Throughout the 1940–1941 period there was continual confusion in discussions of potential supply and demand for steel as a result of these diverse concepts.

The language of basic materials was by no means fixed in definitions pertaining to broad categories of basic materials such as the distinction between carbon steel and alloy steel, between alloy steel and stainless steel, and among other grades of steel, non-ferrous metals, and non-ferrous minerals. Thus, for example, alloy steel, as the steelmaker thought of it, started at the point where the carbon content was roughly less than .5 of 1 per cent and the manganese, silicon, copper or other alloy content was about 1.65 per cent. For many consumers, however, such low alloy figures were insufficient to meet their concepts of alloy steel. Stainless steel was distinguished from alloy steel by consumers but was generally considered alloy by steel makers. Actually, many people in the steel industry loosely defined as carbon steel a sub-

stantial part of the types of steel which non-technicians regarded as alloy. The reason was that low alloy content, particularly manganese, had always been recognized by steel producers as higher grades of carbon rather than alloy. Among steelmakers themselves the distinction between carbon and alloy steel was not always clear-cut because it depended on processing techniques as well as chemical content. In many cases, what was carbon to one mill was alloy to another. And peacetime concepts were modified to fit wartime specifications, further blurring the distinction in common usage.

Even more confusing was the nomenclature of basic material finished shapes and forms. Once again, difficulty sprang from varying popular usage in industry and a lack of familiarity with established nomenclature on the part of those whose experience in industry was limited. In the steel trade, for example, any finished mill shape more than 3/32 of an inch in thickness and 47 inches or more wide was considered a plate and anything under these dimensions was a sheet. To a manufacturer or a steel barrel or drum, however, all drum stock might be considered sheet whether or not it was over or under these limits. Some shipbuilders ordered "plates" when specifications carried a weight of over 25.5 lbs., per square foot. To some consumers, plates were all sizes over 7.65 lbs., per square foot and over 48″ wide; to others, sizes over 10.2 lbs., per square foot; and so on. On other steel shapes also, such as rod and bar, structural shapes, pipe and tube; in copper shapes; in aluminum shapes; and in other materials; usage was far from uniform. What was one man's sheet was another man's plate; what was one man's rod was another man's bar; what was one man's pipe was another man's tube.

These differences were reconciled under the early M Orders by requiring each producer or consumer of the material to report information in the terms specified on the form, irrespective of the manner in which records were kept. Beyond a particular allocation system, however, definitions and classification remained confused. A manufacturer who had to report comparable information to several product administrators found that each used different definitions and classifications. A producer of welding electrodes, for example, found that in reporting his consumption of alloying elements the definitions of tungsten, vanadium and molybdenum alloys changed with each person with whom he dealt.

Such problems not only created difficulties for manufacturers but raised obstacles at many points in the administrative process. Product administrators found it difficult to unify their treatment among differ-

ent manufacturers using the same product, a situation which was reflected not only in their administration but also in top policy. In reports of administrators to top staff, separate product administrators often came to startlingly different conclusions. It was often difficult for policymakers to know whether one report was more accurate than another, whether one report used poorer figures than another, or whether the conclusions of various reports were really identical but expressed differently because of differences in terminology.

Complications increased as questionnaires for each material under allocation demanded greater detail. The tighter supply became in relation to demand the more detail administrators required. Under M-1-i, for example, aluminum producers were asked to report on Form PD-8 only their primary and secondary monthly shipments by nine preference rating categories. This form was promptly replaced by PD-26 upon which aluminum producers, refineries and smelters were asked to report on twenty specific aluminum shapes and forms. Through continuous expansion of reproting requirements aluminum producers were finally asked to report on CMP-24 a total of 43 shapes and forms. Eventually a listing of 77 aluminum shapes and forms was developed but the number for which reporting was required did not go beyond 43. Comparable expansion in other basic materials such as steel, nickel, copper, magnesium, cork and rubber also took place.

In September 1941, a movement began for the compilation of a standard materials list which would define precisely the meaning of basic materials, grades, shapes and forms and other species of the same genus. By this time the work of the Bureau of the Budget, noted above, had resulted in the publication of the *Standard Materials Classification.* A new materials list was developed from this volume and became known as *Materials List #1.* It embraced all of the important metal shapes and forms and included in varying detail such important commodities as textiles, lumber, miscellaneous minerals, and chemicals.

When this list was published the Army made known that it also was preparing a standard materials list. When the Army and the OPM lists were compared many inconsistencies were revealed. By means of compromise, modifications, and adjustments, the official "Red Book" or *Official Classification List of Raw and Basic Industrial Materials* was issued in February 1942 as a joint publication of the ANMB, the WPB, the Navy, and the Maritime Commission.

Crystallization of common usage by the publication of the "Red Book" solved the basic problems of material terminology. Thereafter the battle was more one of improvement than of struggle with basic

concepts. Nevertheless, the use of material shape and form classifications on various reporting instruments varied widely. Items on *Materials List No. 1* were used for the first time on form PD-25A. Throughout the year 1942 and in the early part of 1943, PD-25A continued to incorporate this classification. In the meantime, various questionnaires used in connection with M Orders, inventory controls and P Orders continued their expansion of material shape and form reporting requirements. Since the use of neither *Materials List No. 1* or the "Red Book" classification was compulsory, material breakdowns on these questionnaires often varied considerably from the standard listings and PD-25A.

The movement towards an expanding list of material shapes and forms for copper, aluminum and steel allocation controls was cut short with the introduction of the CMP. For allotment purposes the CMP materials listing was short. The 85 shapes of carbon and alloy steel used on PD-25A were reduced to two CMP groups; carbon steel and alloy steel. The 24 copper and copper-base alloy shapes and forms of the PRP were reduced to four under the CMP, and the 21 aluminum shapes were lumped into 8 shapes and forms.

Other materials allocated by the PRP such as magnesium, zinc, nickel, and lead were not included under the CMP. But by the time the CMP was introduced their classification had become more or less standard. Condensation of shapes, and forms, such as took place under the CMP, did not occur with most other materials under individual allocation programs.

The decision to telescope basic material shapes and forms resulted from a combination of circumstances. To begin with, an expanded list of shapes and forms such as that used under the PRP would have made CMP completely unworkable, particularly in light of the first proposal to issue allotments on a monthly basis, to collect detailed bills of materials for all products and programs, and to ticket each allotment with a document as it filtered through industrial processes. Second, there was a natural revulsion from the lengthy listings included in PD-25A which became cumbersome both from the standpoint of manufacturer response and WPB administration. Third, experience with previous metal-allocation systems had demonstrated that the WPB's control, at least as far as authorizations were concerned, could be just as effective on the basis of broad material groupings as on that of detailed shape and form breakdowns. Individual controls could be used to resolve mill production and distribution problems for tight shapes and forms.

Not all the CMP operations were in terms of the telescoped groups noted above. In the original draft of the CMP there was established material shapes and forms under each main controlled material as follows: carbon steel, 13; alloy steel, 10; copper, 9; and aluminum, 33. Some CMP and related forms required reporting in variations of these shapes. CMP-26 (steel producers' daily reports of accepted authorized controlled material orders for carbon and alloy steel), for example, listed 46 carbon and 33 alloy steel shapes and forms for reporting purposes. CMP-22 (producers' report of shipments of steel castings) required data concerning 10 grades of castings. Aluminum suppliers reported shipments and unfilled orders on CMP-24 by 43 shapes and forms, and cumulative authorized orders accepted for delivery on CMP 12 by 22 shapes and forms. For copper the same sort of situation existed. On form WPB-3112, (brass mill application for copper, copper-base alloy, ingot and scrap) reporting was in terms of 10 different types and on WPB-3508 (brass mill program report) the respondents reported in terms of 4 brass mill shapes. These reporting instruments were typical of a large number of forms pertaining to mill shipments, applications and requirements.

There is considerable doubt whether telescoping of shapes and forms under the CMP was completely desirable. While it is true that summarization of shapes for the three controlled materials eliminated a substantial amount of posting, the absence of detailed shape and form information led to rules and regulations which might not have been necessary with more basic factual detail. Condensation under the CMP was partly based on the assumption that shape and form details would be available for products on prototype bills of materials. But few accurate bills of materials were ever received. No particular problems were created so long as fabricating facilities were balanced with the requirements of end programs and mill capacity was sufficient to satisfy the really essential demands. Under the operation of the CMP, however, there was a constant shifting of the degree of tightness among various material shapes and forms, in and out of detailed sizes, and in various shape qualities and characteristics. This experience necessitated the introduction of new methods to plug the various holes when they opened. The net gain, therefore, was not so great as might appear at first.

General Terminology

Precise meanings of many terms and phrases had to be clearly stated in each questionnaire to assure uniform responses. The problem of

nomenclature, therefore, was more or less as broad as the complete lexicon of the spoken language.

WPB questionnaires showed continual improvement from loose and ill-conceived instructions on many early questionnaires to quite clear-cut and precise instructions on later questionnaires, oftentimes reinforced by reference to well-defined regulations and orders. But no matter how carefully instructions were written, they were always subject to criticism, misinterpretation and misunderstanding. On the one hand were those who objected to the "legalistic" and precise terminology used, and on the other hand were those who deplored the looseness of phraseology. Some manufacturers objected to the insufficient use of technical and "shop language" whereas others complained of phrasing which was subject to misinterpretation. Our interest here is centered largely in a recognition of such problems together with a few typical illustrations of points of difficulty, rather than in a definitive analysis of either the problem or of its solution.

The term "capacity" was continually misunderstood on WPB reports. On WPB-3002 (manufacturers monthly report of unfilled orders and schedule of shipments) and WPB-2578 (report on speed reduction units), for example, manufacturers were required to show in units their productive capacity and schedule of deliveries for each type and size of product. Producers of electrical testing equipment, resistors, electric motors and generators, speed reduction units and unmounted gears inveighed against the term "capacity." They insisted that since two or more models could be manufactured with the same facilities and the same wage earners, the productive capacity for any given size and type depended upon the facilities in use at any given time, and that one month's operations could require heavier production of one model and lighter production of others, while in the next month the situation could be reversed. Even though instructions often stated that capacity by types should follow the proportions established by past actual production and that all capacity data ought to be concurrent, some manufacturers maintained that productive capacity for each model was something which did not exist in fact. Other manufacturers alleged that the term "capacity" was too indefinite. They were not certain whether it meant theoretical capacity, practical operating capacity under the existing conditions of material, labor and machine tool utilization, or actual capacity based on past performance. Although the definition of the term was apparently clear to those in the WPB responsible for its use (the term generally meant the maximum monthly rate of production of each item which could be reached and maintained with the

equipment available, assuming the availability of materials and the maintenance of maximum hours of operation) to industry the term lacked meaning.

The phrase "delivery schedule" raised considerable confusion. WPB-1801 (production of electric motors and generators), for example, requested for each model whose production was reported the standard delivery time in weeks. Industry replied that there was no such thing as a "standard delivery time" for a particular product. To have significance to a manufacturer, the phrase "delivery schedule" had to relate to shipments, since a manufacturer scheduled his production against dates of shipment from his plant and not dates when delivery was made to a customer. Closely connected with these words was the term "requested for delivery." WPB-3002 and WPB-2578 (noted above), for example, called for units of product or dollar value scheculed for shipment and requested for delivery by months. Instructions asked the respondents to specify the total dollar value or number of units included in the total unfilled orders broken down by the amounts that purchasers requested in each of the indicated time periods. Industry insisted, however, that it could not develop the data from existing records, nor could it see how the data could be developed. This resulted from industry's interpretation of the data request to mean the amounts originally requested by customers before placing their orders, information which of course was unobtainable. The intention of the WPB was to compare the items scheduled for shipment with those requested for delivery so as to be in a position to determine the extent to which producers were falling behind schedule.

These illustrations are no more than indicative of a problem which baffled both the WPB and industry throughout the war period. There were hundreds of other words and phrases which caused considerable difficulty, such as idle, excessive, obsolete, usable and nonusable inventory; net shipments; principal product; captive operations; maintenance, repair, and operating supplies; repair parts; and so on. Experience demonstrated that a considerable period of time was often required before words and phrases were so defined that they were mutually comprehensible by the WPB and industry. It also left little doubt that for best results terminology used in questionnaires had to conform as closely as possible to that generally used in industry and that the more precisely a term or phrase was defined the less opportunity existed for misinterpretation and consequent inconsistency in data reported by manufacturers and used by administrators. But what conformed to industrial usage and what was precise were determined only by long and hard experience.

Industrial Reporting Units

An important, but often overlooked question was that of the precise area to be covered by a single questionnaire. One of the basic common denominators determining the usefulness of collected information was the industrial unit for which reporting was to be made. The basic common denominators in practically every industrial statistical report of the WPB were: (1) individual products (or materials), (2) individual industrial plants, and (3) companies. The structure of industry, and in turn the subject matter of production and distribution control, demanded that this be so.

The fundamental distinction between a product (or material) report and a plant report was simple. All questions in a product or material report related to specific products or materials; individual products or materials constituted the subject of the report. The plant report on the other hand used an industrial establishment in one geographical location, or a subdivision of the plant, as the broad subject and required information about all products, all materials and other productive elements receiving attention within the plant area. As such, of course, a plant report was an aggregation of product or material reports but covered all products and materials as well as other relevant productive elements of a plant. In some plant reports all reported data about a plant related to individual products made by the plant or individual materials used in the plant. A company report was often a consolidation of either individual product or plant information, seldom the latter, on a company-wide basis irrespective of geographical location of individual plants or the production locale of individual products or materials.

These distinctions between plant and product reports were hazy in the earliest surveys. Fundamental problems confronting the OPM in early 1941 concerned industrial capacity for certain products or materials, inventories, and production and use of a limited number of individual materials and products in short supply. Early questionnaires sought facts about individual products and materials without a conscious distinction between the plant and the product approaches to informational requirements. Reports had to be sent to plants or companies for completion, not to products. If the needed information related to capacity or to inventories the report generally used a plant as the reporting unit, for data relating to such subjects were reportable and had meaning only in terms of a plant. If, on the other hand, the schedule concerned individual products or materials the result usually was a product report which confined questions to an individual product or material. Frequently, however, these reports (OPM-6 and OPM-23), for instance, combined questions as to capacity with data requests for

individual products or materials with the result that the distinction between a product and a plant became blurred and a questionnaire often in fact used both a product and a plant as a basic reporting measure. But as the war progressed the distinction between the two reporting bases became clearer.

Reporting by industry in these fundamental units was seldom uniform from survey to survey for the same questionnaire or different questionnaires. Thus, for example, specific data on one plant report, such as WPB-732, might include a universe slightly different from that reported on another plant report, such as PD-25A. Great variation existed from report to report because of flexibility permitted in reporting and variation in the field of activity covered by a given set of industry records. Even greater variations existed in factual reporting for specific products.

Product and Material Units of Measure

Closely allied to the names and classifications of products and materials were their units of measure. Here again the WPB found widespread differences in reporting of standard quantities and values. And here again the trend was towards greater precision in definition; and more general use of those standards which it was found were mutually understood by both industry and government, readily reportable by industry and acceptable for administrative purposes. Measurement units generally were necessarily fixed according to industry usage and industry records, particularly when they related to past activities. Unfortunately, however, usage was not always standard among plants and industries and usage common in industry was not always of most utility to government.

Material units of measure generally fell into two distinct groups, or else defied classification. The most widely used was a weight measure. Most basic materials, particularly the metals, originally requested reporting in terms of tons or pounds. Even these simple concepts created difficulties. Some questionnaires requested information in long tons, others in net tons, and still others in short tons, generally without defining the standard specifications pertaining to the concept used. Some forms mixed the reporting requirements, such as PD-209A (report of vanadium consumption), and data were requested in both long and short tons with no definitions of either. The result was a 12 per cent variation in reporting. Experience led the WPB to request reporting for most materials, when tonnage was specified, in short tons of 2000

pounds each. Reporting in pounds presented no particular problem because a pound was solidly entrenched as a standard unit.

For those materials which were incorporated in other basic materials at the mill level, particularly the ferroalloys, reporting was often required in terms of the content of these other materials. PD-358 (producers application for molybdenum for melting), for example, requested the per cent of molybdenum in finished products. PD-359 (producers report of molybdenum containing metal) requested reporting by pounds of molybdenum content. Complications developed from time to time in the use of such measures. Nitrogen-containing materials were at first measured in tons of ammonia content, then ammonia equivalent, and finally in tons of calculated nitrogen content.

Units of measure outside these areas were too diverse for ready classification. PD-338 (inventory of jewel bearing material) measured jewel bearing materials, such as natural sapphire and synthetic ruby, in carats. PD-236 (consumption of jewel bearings) established reporting in terms of the number of jewels contained in bearings. Mercury was measured in flasks, phosphates in terms of equivalent phosphorus pentoxide, ethyl alcohol in gallons of 190 proof, penicillin in billions of oxford units, and leather in terms of equivalent hides and skins.

Three main types of units of measure were commonly used in connection with products. The first and most desirable unit for products, and the one most frequently used, was one unit of a multiple of the product in question. This was satisfactory when the product was uniform, such as hot drink cups, telephones or kerosene railroad lanterns, but could not well be used in the case of products presenting diversity in important characteristics. The unit classification was strained in numerous instances in summary tabulations, as in the case of steam and diesel locomotives, scales and balances and various machines. Second, many orders established controls on the basis of the tonnage of iron and steel or other metal used in the production of the controlled products during a base period and reports were made in terms of units, pounds, or value of the finished product. This material content approach, of course, was limited to availability of records and was not as widely used as the other methods. Third, when the use of units and size or weight measures was unsatisfactory, dollar value measurement was used. This was inherently unsatisfactory as a control measure because of lack of information as to quantity and quality of production and distribution, number of units, quantity of raw material and labor required, and changes in prices. It was used in lieu of anything better and at the same time was practical.

Confounding problems often arose because of the use of varying units of measure and the difficulties of settling upon a satisfactory conversion formula to secure comparability of related data. The Foreign Economic Administration, for example, received from the WPB allocations of leather and textiles in units of equivalent hides and linear and square yards and pounds of cloth or yarn, but issued licenses and made requisitions in terms of shoes and finished textiles. Months were required to develop conversion factors and standardize terminology and measurement units. Comparable problems were encountered in many other products and materials.

Difficulties with varying units of measure resulted in a trend in the WPB, continuing through the war years, to require reporting of material and products in terms of individual units such as quantity units for products and tons and pounds for metals. The coincident development of product and material definitions and classifications when related to uniform units of measure made the data collected by the WPB increasingly more useful with each step forward in these areas.

Time Periods Covered on Reports

Under the rapidly changing conditions of a war economy a year was too long a period to wait for reports upon which to base policy, and time periods moved from quarters, to months, to weeks, and often to days depending upon the urgency of a problem. From time to time products were allocated on a daily telephonic basis. Broad material allocation programs, however, could not operate on so short a time period. Many of the individual allocation systems in metals existing up to and through the Production Requirements Plan operated on a monthly allocation basis. The PRP itself was based on quarterly reporting and allocation. This, it was felt was too long for many tight items, and the early promulgation of the CMP envisioned monthly allocation. It was quickly perceived, however, that such a time period would involve a tremendous amount of paper work, conferences and decision writing. As a consequence the time period instituted with the first operation of the plan was the calendar quarter. As circumstances permitted, other allocation plans quickly followed the quarterly allotment procedure, and the calendar quarter finally evolved as the best time unit for making allocation, scheduling production and reporting on implementation.

Measures and Definitions of End-Use Data

The general approaches to production control embodied in the preference rating system, P Orders, and M Orders, encountered a problem

relating to identification of end programs at sub-manufacturing levels which frustrated the WPB throughout the war period. Early operators of material allocation controls recognized that their decisions concerning mill shipment authorizations were not particularly intelligent in the absence of information relating to the uses, particularly the end-uses, to which the basic material ultimately was to be put. This type of information was difficult to acquire because no workable system existed for passing information from the top to the bottom of the industrial system.

There were many different approaches to the development of end-use information. Throughout the year 1941 and into early 1942 constant difficulties with old methods resulted in their revision or the institution of completely new systems, each striving to obtain information about end uses which was considered vital to the operation of each allocation plan. By the spring of 1942, after almost one full year of experience, methods to derive end-use information were little closer to their goal than in the spring of 1941.

The approaches current in the spring of 1942 may be classified into several rough categories. (1) The *general* approach was adopted in PD-69 (consumers pig iron order for shipment) which requested applicants to give a complete description of the ultimate use or product to be made from pig iron for which application was made. With little help, applicants were asked to be very specific, i.e., "Pump castings for Navy ships." (2) PD-123 (producers report on copper and copper alloy deliveries) embodied what may be classified as a *use* approach. This form originally contained 37, and ultimately 100, classifications of end uses for which producers were to report deliveries of various copper and copper-alloy shapes and forms. On the basis of the classification given it was not only impossible for producers to file completely, but one end product conceivably could appear in three or more classifications as set forth, depending upon its ultimate use. As a result, tabulations of this form were often weird and meaningless. (3) The *Product-type* approach may be illustrated by PD-180 (customer application for phenol). The applicant was obliged to enter the name of the product into which he was incorporating phenol and the specific ultimate use of that product. (4) The *use and product* approach was typified by PD-197 (producers report of lead and lead alloy shipments). On this form 33 classes of products were preprinted in the stub, such as storage batteries, cable covering, and white lead. A cross-classification was requested showing the amounts of lead used in each of the preprinted products.

The persistence with which this type of data was sought and the

willingness with which the WPB introduced rapid changes in past methods, each more complicated than its antecedent, testify to the importance attached to this sort of information and the desperation with which it was sought. Neither the material procurement agencies nor the WPB commodity divisions were in a position to match demand-supply relationships on a factual basis. As a result it not only was impossible to assign the resources of the country to those uses of most importance, but it was difficult to determine the desirability of expanding production of critical materials. The key to the puzzle seemed to lie in the development of machinery by which accurate identification of end products manufactured by prime contractors could be transmitted throughout the entire industrial system, through intermediate production and components, to the basic metal mill producers. Such a system, it was reasoned, would allow the WPB to follow the distribution of basic materials and components of each important end-product and assure its completion on a balanced schedule and in conformance with a carefully pre-arranged principle.

To meet these objectives a bold experiment in end-use identification was prepared in June 1942 with the issuance of Priority Regulation No. 10. The regulation observed that its purpose was "(1) to obtain standardization and reduce the rules of different forms for allocation purposes that industry now must submit to the War Production Board, and (2) to furnish information needed by the War Production Board in the allocation of materials. . . . To allocate intelligently," it went on, "it is necessary to know the subdivision of the war, industrial and civilian programs for which materials are going to be used. For example, it is essential to know whether the materials are going to tanks, or destroyers, or railroads, or to office machinery and supplies in order to determine allocation policy. In addition, it is necessary to know in a general way what type of purchasers will eventually receive particular products of industry, i.e., whether the product will eventually be delivered to the Army, Navy, etc."

The theory behind the code was simple. Each purchase order from a prime contractor would include an end-use code indicating the use to which the purchased materials or products were to be put. Each subcontractor would then repeat the code symbol on his orders for materials and parts until, in theory, the symbol appeared on orders placed for basic materials. Suppliers of critical materials would then be in a position to report their shipments by end use codes, which would reveal information on both the end use and end user. It was recognized that this plan would not be feasible in every case, but it was thought

that the band of unidentified end use could be narrowed and that the system would provide the first real check on the relationship between end products and scarce materials.

The WPB materials and commodity Divisions violently protested the use of this system. Having developed end-use reporting requirements to fit their particular needs, they insisted that the classifications in the end use code were so broad as to be almost meaningless. The purpose of this code was to standardize the many existing codes and to keep the number of classifications to a minimum, so as to reduce the burden on industry. Yet if the code did not meet the needs of the Divisions it was doomed to failure. To say this is not to say that the end use codes developed by the Divisions were workable. On the contrary they lacked uniformity and could not be tabulated. But they did provide the narrow bands of end use which the Divisions wanted and which the newly devised code lacked. The theory behind the code may have been sound, but it fell down in its practical applicability.

Serious criticism also came from many industries. The steel industry, for example, complained that it was faced with serious difficulties in the application of the regulation to its normal business practices—so much so that the data received by the WPB from steel mills had little value for allocation purposes. Fundamental problems of the steel industry lay in three areas. First, steel warehouses were required to analyze monthly all their purchase orders to determine the allocation classification pattern which they would use on their own purchase orders which they placed with steel suppliers in the following month. Since the warehouses received thousands of small orders each month, the accounting and tabulating problem was staggering. Moreover, since the Steel Division of the WPB thought it could not and therefore would not allocate steel in accordance with the information furnished by the classification system, all the paper work involved was useless. Second, since the warehouses added extra charges for cutting, handling, etc., analysis of the sales pattern by dollar volume as required by the regulation, was misleading. Third, since the Steel Division of the WPB required information according to the nine purchaser classifications used under its own Order M-21 instead of the five purchaser symbols used under PR-10, the industry had to keep two sets of records.

Following a short but turbulent life the system was abandoned in November 1942 as a failure. It was cumbersome in use, unworkable in large segments of industry, particularly at the component level, and produced unreliable order identification on mill order books. The concept was valuable but the methodology was poor.

Throughout the experience with Priorities Regulation No. 10, it was becoming clear that satisfactory control over war production could not be answered by a suitable end-use system slone. It was but one ingredient in a far broader need which embraced the balancing of end programs among themselves in relationship to the totality of scarce materials; the scheduling of materials through the industrial system to end products, in terms of quantity and time; the adjudication of competing claims at critical levels of fabrication and direction of mill fabricating facilities.

Following the failure of Priorities Regulation No. 10, the Production Requirements Plan provided part of this needed method. The introduction of the Controlled Materials Plan in 1943 furnished a practical mechanism by which end program symbols could be transmitted down through the contractual chain to mill levels for controlled materials. The plan accomplished this simply by requiring that a controlled material allotment and program symbol had to accompany each purchase order if the product ordered used any one of the controlled material shapes and forms. Not only were prime consumers requested to use major program symbols to identify orders carrying controlled material allotments but they also identified rated orders for class B products for which they did not reallot quantities of controlled materials. By means of this procedure end-program identification, based on CMP allotments, was provided for a very large part of American industry, from procurement agencies through to metal mills.

The system by which Claimant Agency symbols were passed down to Class B producers was not designed to provide precise identification of either Claimant Agencies or programs. To begin with, many manufacturers operating under P, U, and other orders and regulations used a variety of specialized symbols not identifiable in terms of Claimant Agencies. Second, rated orders from prime consumers did not always carry Claiment Agency symbols. Finally, extension to Class B producers by prime consumers was in terms of Claimant Agencies themselves, and books were not maintained showing program symbols. The result was a broad specification of end-use on CMP-4B applications. Despite many difficulties in identification growing out of weaknesses in parts of the system, the Controlled Materials Plan solved practically, from an administrative point of view, the end-use problem for controlled materials. The solution, however, was not so much imbedded in a method of classification as in a complete new comprehensive allocation system.

Individual allocation systems continued to exist for materials and products other than those under general comprehensive programs such as CMP, Lumber, and Rubber Tires. The methods for determining end uses were compromises between the desires of the WPB and the records of industry. As might be expected the methods varied considerably and incorporated the same features which were used under M Orders. As a consequence they engendered the same pre-CMP problems of compliance, the same type of industry criticism and the same collection of inadequate data. For many materials and products the problem was never solved satisfactorily. The discipline created by Priorities Regulation No. 10 and the operation of the CMP helped, but industrial practices negated passage of complete end-use information to raw material sources without a superimposed method such as the CMP allotment chain.

The length of time required by the WPB to develop satisfactory definitions and classifications should not be viewed too critically. Of course the struggle was much influenced by the fumbling which characterized the entire evolution of production and distribution controls. But the matted problems of definitions and classifications were, even beyond that, hard to disentangle. Even when definitions were clear cut new circumstances inevitably arose to blur them. In some fundamental aspects, the problem of nomenclature specification was insoluble. What was a sound definition depended upon the persons using it, industry usage, the state of classification systems and the precise needs of a given control policy in which it was a part. The broad definition of a material such as carbon steel, to cite one instance, was reasonably satisfactory in the early war period. But it really meant little. To speak of carbon steel sheet helped a little, but not much. Further clarification of carbon steel sheet in terms of a specific gauge, size and shape was better, but still not sufficient for complete clarity. To carry the needed extensions of definition to extremes, which would be necessary for complete and unequivocable understanding, obviously would have entailed the most minute description of characteristics of a product or material. Even this would not be enough. In a war period, rapid changes take place in old products and there is a constant introduction of new products with new names, characteristics and uses. In such circumstances workable definitions and classifications are all that can be expected. Much time and experience, however, are required before this condition can be reached.

[1] *Standard Industrial Classification*, Central Statistical Board (through 1937); Bureau of the Budget (1940 to present). Part I, Manufacturing Industries, December 10, 1937. (Tentative draft.)

Volume I, Part I, Manufacturing Industries, List of Industries, May 10, 1938.

Volume I, Part I, Manufacturing Industries, List of Industries, 1939.

Volume I, Part III, Manufacturing Industries, Alphabetic Index of Products, Establishments and Processes, 1939.

Volume I, Part IV, Manufacturing Industries, Alphabetic Index of Products, Establishments and Processes by Major Industry Groups, 1939.

Volume II, Part I, Non-manufacturing Industries, List of Industries, 1939.

Volume II, Part II, Non-manufacturing Industries, Description of Industries, 1940.

Volume II, Part III, Non-manufacturing Industries, Alphabetic Index of Products, Establishments and Services, 1940.

Volume II, Part IV, Non-manufacturing Industries, Alphabetic Index of Products, Establishments and Services by Major Industry Group., 1940.

Volume I, Part II, Manufacturing Industries, Description of Industries, 1940.

Volume I, Manufacturing Industries, 1941.

Volume II, Non-manufacturing Industries, 1942 (still current).

Volume I, Part I, Manufacturing Industries, Titles and Descriptions of Industries, November 1945.

Volume I, Part II, Manufacturing Industries, Alphabetic Index, December 1945.

Volume II, Non-manufacturing Industries, 1942, available at GPO for 25¢.

Volume I, Part I, November 1945, available at GPO at 75¢.

Volume I, Part II, December 1945, available at GPO at $1.00.

Standard Classified List of Commodities, Prepared under auspices Executive Office of the President, Bureau of the Budget; War Production Board; and Department of the Treasury, Division of Procurement. First edition (for comment only), July 1942. Reg. First edition, December 1942. Revised edition (Vol. 1), May 1943. Vol. II of Alphabetic Index, February 1944. Supplement to Vol. II, November 1945. July and December 1942 reports not available except in the files of the Bureau of the Budget. Other reports available at U. S. Government Printing Office (May and Feb. 1943 reports @ $1.25, and November report @ $2.00.)

BOOK REVIEWS

Edited by

OSCAR KRISEN BUROS

Rutgers University

The Teaching of Statistics: A Report of the Institute of Mathematical Statistics Committee on the Teaching of Statistics. *Harold Hotelling*, Chairman (Professor of Mathematical Statistics and Head of the Department; Associate Director, Institute of Statistics; University of North Carolina, Chapel Hill, N. C.), *Walter Bartky* (Dean of the Division of the Physical Sciences and Professor of Applied Mathematics, University of Chicago, Chicago, Ill.), *W. Edwards Deming* (Adviser in Sampling, Bureau of the Budget, Washington, D. C.), *Milton Friedman* (Associate Professor of Economics, University of Chicago, Chicago, Ill.), and *Paul Hoel* (Associate Professor of Mathematics, University of California, Los Angeles, Calif.). Reprinted from *Ann Math Stat* 19(1):95–115 Je '47. Washington 25, D. C.: National Research Council (2101 Constitution Ave.), 1948. Pp. 95–115. Paper. $0.25.

The Teaching of Statistics in Universities and Colleges. Based upon a report by the Royal Statistical Society's Committee on the teaching of statistics: *E. S. Pearson* (Chairman), *R. G. D. Allen, H. Campion, William Elderton, C. Oswald George, R. F. George, M. Greenwood, D. Heron, J. O. Irwin, H. Leak, E.C. Rhodes, R. Stone, L. H. C. Tippett*, and *J. Wishart*. Preprint of *J Royal Stat Soc* 110(1):51–7 pt 1 '47. London W.C.2: Royal Statistical Society (4 Portugal St.), 1947. Pp. 7. Paper. 6*d*.

REVIEW BY ALAN E. TRELOAR

Professor of Biostatistics, University of Minnesota

Minneapolis 14, Minnesota

SHORTAGE of qualified personnel within the statistical profession has been a matter of concern to employers demanding services from that source. The reports reviewed herein clearly testify that those charged with leadership in recruiting and training such personnel not only share that concern, but are acutely aware of grave problems basic to the unbalance which, if not removed, will lead to continued deterioration as well as wastage of educational opportunity in a very broad sense.

Appreciation of the usefulness of statistical methods was limited to a few farsighted pioneers at the beginning of this century. Response to their demonstrations of new power in analysis was very slow in developing. Biologists and workers in the social sciences generally were not able to assimilate the new biometry being created for them; for the most part they not only lacked the mathematical training so necessary to appreciation of its approach, but were most reluctant to believe that there was any hope that the variations which beset them were amenable to mathematical treatment. Only a few of broader vision saw great hope in the new tools and applied themselves to work on the frontier with tremendous enthusiasm despite their poverty in mathematics. With indefatigable energy they set an example in

both teaching and research that gradually leavened the main lump. Their teaching was eminently practical. It had to be, for mathematical abstraction quickly bereft them of followers. Students learned by analogy how to apply the methods rather than to comprehend fully the basic principles of a new philosophy and the limitations imposed by the simplifying assumptions made by the theorists to reach the first solutions. These students have in turn become teachers, carrying responsibilities for which they were equipped only in a very limited sense.

In reviewing the present situation it is well not to lose sight of the contributions made by these pioneers with all their handicaps. Statistics enjoys a large demand today and is what it is largely because of them. Theoretical statisticians who now strive to serve the sciences and assume leadership responsibilities for the future development of statistics should properly temper criticism, no matter how justified it may seem from specific points of view.

The reports being reviewed herein, differing sharply in their emphasis on criticism, are in close agreement when considering what should be done in the future. Both reports were preceded by another* from the Committee on Applied Mathematical Statistics (L. P. Eisenhart, S. S. Wilks, C. I. Bliss, E. U. Condon, H. O. Gulliksen, L. J. Reed, C. F. Roos, W. A. Shewhart, H. M. Smallwood, F. F. Stephan) of the National Research Council, which urged the same objectives. All three should be considered together. The American point of view is thus more adequately represented, both in numbers and in balance of practical and theoretical components. The Royal Statistical Society report is based on the deliberations of many persons representing a wide coverage of both theoretical and applied statistics.

One need not review here evidence submitted of interest in and need for statistical methodology. A brief summary of the proposals for meeting the need would be in order, however. An introductory course of an orientation type is recommended for all college students as an essential component in a truly liberal education. Probability has replaced certainty in the modern way of looking at things, and those not familiar with the concept of aggregates and the fundamentals of the statistical argument in relation to it face a difficult mental reorientation in their efforts to keep abreast of life's affairs. The committees look forward to the day when this orientation will be a regular part of the secondary school education, but for the immediate future provision must be made for it to be given in the college curriculum. It is judged that the course would most profitably be built around experimental demonstration; it should not require special mathematical ability.

A basic course in statistical reasoning and methodology covering principles common to all applications and given centrally for all students needing more than just a general orientation is strongly urged. The present system

---

* *Personnel and Training Problems Created by the Recent Growth of Applied Statistics in the United States.* Reprint and Circular Series Number 128. Washington, D. C.: National Research Council, May 1947. Pp. v, 17. Paper. $0.25. Reviewed in this JOURNAL (Dec. 1947), by George W. Snedecor and Holbrook Working.

of many departmental courses of all qualities is judged to be badly in need of overhaul; applications departments should yield the common ground, concentrating on facets peculiar to their own special fields, thus fostering interest and competence. It is acknowledged that the general course would have to be given for some time with two levels of mathematical prerequisites, but these would ultimately fuse with calculus as a prerequisite for all. The two courses given meantime should correspond in every way possible and have the same coverage.

Departments of application should continue to give statistical courses reconstituted to accord with the plan stated in the preceding paragraph. These courses might be regarded as intermediate in grade. They should embrace discussions of sources of data, limitations peculiar to them, and applications of methods with special emphasis on practical considerations. Attention would be given here to procedures of interest to the field of application but not of sufficient general importance to be included in the central introductory course.

Offerings at the advanced level would be planned in accord with facilities made available to meet existing needs. The IMS report attaches great importance to the training of future teachers through the route of advanced courses in statistical theory demanding advanced mathematics as a prerequisite. While all reports emphasize the importance of experience in application as an essential concomittant in statistical training at all levels, there is a note of warning in one (IMS) against allowing the specialist in statistical method to participate too extensively in applications lest they "so engross his attention that he cannot keep up with the progress of statistical method itself." History seems to this reviewer to indicate a drift toward theory by those of considerable competency in both theory and applications.

Centralized teaching of statistical method within any college, as envisioned by these committees, calls for creation of a special teaching unit whose personnel would be free to study statistics for its own sake. "In statistics, as in other subjects, a profession of general practitioners requires the support of an inner circle of specialists, research workers and teachers; and the inner circle maintains its strength partly through teaching and serving the general profession. . . . Where this advanced teaching can be associated with an important research school of statistics, considerable advantage will follow" (RSS). An independent department or institute of statistics is favored by all three reports, but where attachment to an existing department is insisted upon administratively the RSS committee recommends affiliation with mathematics. It is preferred that this unit have its own budget and be equipped as completely as possible with all facilities for practical work and research, including easy access to a fully adequate library of original sources and mathematical tables. This unit and the statistical units within applications departments should be drawn into close association through a coordinating committee, each gaining strength through reciprocal aid. Such a committee drawn solely from applications departments must not be accepted as a sub-

stitute for an administratively independent unit devoted primarily to teaching and research in statistical methodology.

The RSS report makes no comment on current standards of teaching in statistics. The NRC committee acknowledges as undesirable the wide variation in quality of statistics courses offered by applications departments within the colleges of this country, and the vested interests which the departments have developed in those courses, but goes no further. Only the IMS Committee has the courage (if courage it be) to attack forthrightly the procedures of appointment and administration of statistical personnel within the applications departments. An olive branch is offered in statements of the form: "All this does not mean that any considerable number of people teaching statistics are unworthy or objectionable members of the academic community. Many, indeed, are of superior intellect, upright character, personal charm, and undoubted teaching ability. Some are making creative contributions to other subjects. The only trouble is that they are teaching a subject in which they are not specialists, and which progresses so fast that only specialists can keep up with it." The dove of peace, laboring somewhat in its flight, makes no call at the offices of the responsible administrators. The argument is plausible, skillfully phrased, and reading it provides some entertainment. But this reviewer doubts that it will add to attainment of the objectives with which the report is so much concerned. A supply of adequately trained personnel in the sense defined has not been available. Theoreticians were very slow in responding to the need they are now so anxious to fill. They will be welcomed with unquestioning enthusiasm just as soon as they demonstrate that they can do the job more effectively than their predecessors. The affinity of theoretical statistics and mathematics is unquestioned. The report admits that mathematicians have so behaved in their teaching as to make children "resolve to quit mathematics as soon as possible." Nothing surely would be more welcome than clear demonstration that such performances will have no place in the techniques adopted by theoretical statisticians in undertaking the heavy responsibility of teaching statistics on a college wide scale.

To say these reports are important is to belabor the obvious. They constitute a most commendable attempt to plan for a better future. All teachers of statistics should read them carefully and reflect constructively on the problems they attack.

Investment, Location, and Size of Plant: A Realistic Inquiry Into the Structure of British and American Industries. *P. Sargant Florence* (Professor of Commerce); assisted by *W. Baldamus* (University Research Fellow). (University of Birmingham, Birmingham, England.) National Institute of Economic and Social Research, Economic and Social Studies VII. London N.W.1: Cambridge University Press (Bentley House, 200 Euston Road), 1948. Pp. xiii, 211. 18*s*. (New York 11: Macmillan Co. [60 Fifth Ave.] $5.00.)

REVIEW BY EDGAR M. HOOVER
Staff Economist, Council of Economic Advisers
Executive Office of the President, Washington 25, D. C.

PROFESSOR Florence reports here on an analysis of three interrelated attributes of industries: prevailing size of establishment, degree of geographic concentration, and invested capital per man. The basic data are primarily British and American production censuses of the latest available prewar years. He develops measures for each of the three attributes, seeks to explain the observed associations among them, and evaluates the significance of the findings for purposes of public policy.

The book as a whole represents a notable advancement of the state of our knowledge of industrial structure. Its value is enhanced by the presentation of industry-by-industry tabulations of nearly all the statistics discussed.

At times readability and directness appear sacrificed to wealth of suggestive detail. For example, some pages of text and one whole appendix are devoted to measures of the "heaviness" of industries, though the concept of heaviness is not used anywhere else so far as this reviewer could discover. On the whole, the reader will be grateful for the wealth of suggestive loose ends tagged for further exploration.

Size of plant is measured (after a discussion of alternatives) by number of employees. The extreme skewness of the size distribution of plants in an industry is circumvented by using frequency tabulations in which each class limit is 2 or $2\frac{1}{2}$ times the next lower one. All the author claims in behalf of this inelegant procedure is that it seems to work: a distinct modal or "prevalent" size class appears in most industries except where wide variety of sizes is to be expected on a priori grounds.

The aspect of location with which this book deals is the deviation of the geographic distribution of a given industry from that of industry as a whole. The "coefficient of localization" has already gained fairly general acceptance as a result of its previous use by Florence and others, but it is here tabulated for the first time for a nearly complete list of industries of both countries.

The "intensity of investment" of an industry poses the most serious problems of measurement. Horsepower capacity per worker is chosen as the least objectionable of several possible indicators.

When specific British industries are compared with their American counterparts, a striking degree of international correspondence appears in plant size, location pattern, and use of power. This leads the author to his basic theses: that the basic technical operations of industries (and not merely

the national political, social and financial milieu) are significant in accounting for interindustry differences, and that size, location, and use of power are interrelated characteristics.

In analyzing the relation between an industry's typical size of plant and its coefficient of localization, Florence uses simple contingency tables. The observed relation of plant size to localization is, however, as earlier studies by Florence had indicated, a complex one. It goes considerably beyond the range of statistical proof as yet.

Plants in many of the most scattered or "ubiquitous" industries show no definite central tendency in size. In areas of dense concentrated market or material supply, plants can be larger than elsewhere, so a diversity of plant sizes is naturally found within a single industry.

In industries with intermediate degrees of localization, greater localization is associated with larger plant size.

The most highly localized industries are characterized by medium-sized plants because (*a*) in these industries transport costs do not limit size of plants, so plants need not be small to be close to scattered suppliers or markets; and (*b*) close juxtaposition of plants makes it possible to secure many of the economies of large-scale production by virtue of the size of the production center rather than of the individual plant—plants therefore can specialize more narrowly, and do not have to be large to be efficient. This hypothesis as to highly localized industries is tested by a more detailed analysis of concentrations of metal trades in Birmingham and textiles in Lancashire.

In exploring the relationship between size of plant and horsepower capacity per worker, Florence uses the same technique of contingency tables. In general, larger plants are associated with greater use of power per man, greater financial requirements, and greater prevalence of multi-plant firms.

The last part of the book speculates on implications of a continued trend toward larger plants and larger firms. There is little novelty in the author's case that protection of the interests of labor and the consumer will require a better and broader science of industrial management along with watchful public scrutiny or control of monopolistic industries. The most interesting point in the discussion of policy problems is an explanation of the difficulty of economically encouraging small business and dispersion of industry to small nonsuburban communities. According to the relationships shown by Florence's statistical analysis, most industries that are not tied by high transport costs to local markets or sources of materials are either highly clustered or characterized by large plants. Thus the promotion of small decentralized business faces the problem of increasing the efficiency of small plants or breaking up or transplanting complex localized groupings of plants. There may be further conflict between the aims of decentralization in nonsuburban small communities and local diversification of production.

These barriers are not necessarily impassible. The observed associations

of size, location, and investment on which Florence bases his case are sufficiently loose to provide many exceptional industries and plants, and the economic advantages of size and localization are not immutable. Nonetheless the basic difficulties remain and this book has done a useful service by bringing them to light.

Foundations of Economic Analysis. *Paul Anthony Samuelson* (Professor of Economics, Massachusetts Institute of Technology, Cambridge, Mass.). Harvard Economic Studies, Vol. 80. Cambridge 38, Mass.: Harvard University Press (38 Quincy Street), 1947. Pp. xii, 447. $7.50. (London E.C.4: Oxford University Press [Amen House, Warwick Square], 1948. 42*s*.)

REVIEW BY GEORGE J. STIGLER
Professor of Economics, Faculty of Political Science
Columbia University, New York 27

PROFESSOR SAMUELSON'S distinguished doctoral dissertation was largely completed about a decade before its publication (p. vii). This delay in appearance is doubly unfortunate: unfortunate for other economists, deprived of portions of its brilliant theoretical analysis and criticism not hitherto published; unfortunate for Samuelson, much of whose subject matter is currently of much diminished interest. But the problems he deals with have been and will continue to be important and his treatment of them should be read by all mathematically trained economists interested in formal theory.

The leitmotiv of Samuelson's treatise is the search for meaningful theorems, that is, theorems which are in principle capable of empirical contradiction. The statement that a person seeks to maximize utility is (in many versions) a tautology: it is impossible to conceive of an observational phenomenon that contradicts it. The statement that demand curves have negative slopes on the other hand, is capable of empirical contradiction, and hence meaningful. There is no question but that this criterion is an excellent scientific sieve, whereby tautological and observationally undefined theorems (with which economics is well supplied) may be separated out from the theorems with content.

But it is not true that theorems about unobservable things are necessarily meaningless in a social science. On introspective grounds—and this is a perfectly valid and indeed indispensable source of information—one believes that there is such a thing as utility, and that men strive for it. One can dededuce additional conclusions (of a welfare nature) from this information, that cannot be reached by a strict positivist. Thus, we know and find it significant that, with given prices and money income, a consumer feels better off making his own choices than having someone else make them for him. The whole purpose of utility analysis in economics is to permit such welfare propositions, and such propositions are meaningful even though based upon introspection and communication.

Two different, but related, methods are employed by Samuelson to derive meaningful theorems. The first method is by use of the secondary (or sufficiency) conditions for an extremum; we may illustrate it by a very simple example. Let a competitive firm be operating at a maximum profit output where marginal cost equals price; the secondary condition is that marginal cost is increasing at this output. Let the price now rise: if the theory is correct, the output of the firm will rise or at least not fall (even if the marginal cost curve is vertical)—hence if a decline in output follows a rise in price, the theory is contradicted.

Samuelson gives a variety of illustrations of this approach (Chaps. 2–3) and then proceeds to more detailed studies of the output of the individual firm (Chap. 4) and the demand of the individual consumer (especially Chap. 5; special mention should also be made of the excellent discussion of index numbers in Chap. 6). Many of the theorems are already familiar, through the work of Slutsky, Hicks and Allen, Stackelberg, Mosak, etc., but nowhere else can one find so comprehensive a treatment.

There are only two points I would raise in connection with this first method of deriving meaningful theorems. The first is: why should not theorems on groups (industries, for example) also be derived? It is implied (p. 5) that this is impossible because an industry equilibrium does not represent a maximum position, but this is surely mistaken: the particular entrepreneurs are in the industry because it is the most profitable place to be.

My second comment is connected with the question: what if the theorem is contradicted by observation? Samuelson says it would not matter much in the case of utility theory (p. 117); I would say that it would not make the slightest difference. For there is a free variable in his system: the tastes of consumers. These tastes are not observationally (or operationally) defined, so any contradiction of a theorem derived from utility theory can always be attributed to a change of tastes, rather than to an error in the postulates or logic of the theory. The situation is less acute in the theory of the firm, for the corresponding variable (technology) is not observationally so free. But Samuelson's (and everyone's else) utility theory is still tautological as an explanation of market behavior.

Samuelson's second method of deriving meaningful theorems is to draw the consequences of dynamic stability. For example, we may assume that price will rise (fall) if the quantity demanded at the present price exceeds (is less than) the quantity supplied, and that the greater the rate of change of price, the greater the discrepancy between quantity demanded and quantity supplied (p. 263). Then if, after an initial disturbance, the price is to approach an equilibrium value, the demand curve must cut the supply curve from above. A miscellany of examples of this technique are given (Chap. 9), the most interesting being those for a simplified Keynesian system.

Thereafter Samuelson enters on a roving discussion of dynamic theory. Most discussions of economic dynamics, I feel, would better be entitled, "What I Know about Differential and Difference Equations"; with Samuel-

son's mathematical erudition, no doubt his version is, "What the Reader Should Know of Differential and Difference Equations." The economics, as usual, is appallingly absent. Some of the infinities of mathematical possibilities are discussed, but only in the most formal terms; there is no instance of the enlargement of our knowledge of economic processes in our society. Samuelson may reply that he is only providing tools, but who can know what tools we need unless he knows the material on which they will be used?

Although it is not an integral part of the search for meaningful theorems, welfare economics of the Hotelling variety (i.e., somewhat free of value judgments) also receives considerable attention (Chap. 8). I persist in my belief that this is one of the less fertile areas that modern economists till; it is symptomatic that we have elaborate instructions on how to form welfare judgments that do not depend on value judgments, but we have no illustration of the application of this technique to a real problem of contemporary policy. Samuelson indeed offers much support for this skeptical view, by this enumeration of the assumptions of the new welfare economics (pp. 222ff.), most of which are held to be partly invalid. But he fails to examine other facets of the problem, one of which, I think, is especially significant. When economists are writing freely on desirable policy, that is, when they are not writing on methodology, the disputes are almost always over how the economic system works, and not over the goals that should be sought. Thus, although I disagree with some of the policy recommendations in Samuelson's recent textbook, I do not differ (except in emphasis) from the goals that they are intended to subserve.

Samuelson's exposition is consistently mathematical, and presupposes a knowledge of advanced calculus, higher algebra, and differential equations (in appendices he summarizes some theorems from these branches and gives a lengthy and technical introduction to difference equations, which to my unmathematical mind seems much more difficult than say Jordan's book on finite differences). I consider this failure to provide translations for the "literary" economist a serious shortcoming of his work. He dismisses translations into words as "mental gymnastics of a peculiarly depraved type." I disagree. There is no depravity, nor is there virtue, in telling other competent economists things in a language they all can understand—there is simply responsibility to the canons of scholarship.

Sampling Inspection: Principles, Procedures, and Tables for Single, Double, and Sequential Sampling in Acceptance Inspection and Quality Control Based on Percent Defective. By the Statistical Research Group, Columbia University, Applied Mathematics Panel, Office of Scientific Research and Development. Edited by *H. A. Freeman* (Associate Professor of Statistics, Massachusetts Institute of Technology, Cambridge, Mass.); *Milton Friedman* (Associate Professor of Economics, University of Chicago, Chicago 37, Ill.); *Frederick Mosteller* (Associate Professor of Mathematical Statistics, Department of Social Relations, Harvard University, Cambridge, Mass.); and *W. Allen Wallis* (Professor of Sta-

tistics and Business Economics, University of Chicago, Chicago 37, Ill.). New York 18: McGraw-Hill Book Co., Inc. (330 West 42nd St.), 1948. Pp. xx, 395. $5.25. (London W.C.2: McGraw-Hill Publishing Co., Ltd. [Aldwych House, Aldwych]. 31*s.* 6*d.*) *Two reviews follow:*

REVIEW BY EDWIN G. OLDS
Associate Professor of Mathematics, Carnegie Institute of Technology
Pittsburgh 13, Pennsylvania

SAMPLING inspection is the process of deciding whether to accept or reject an aggregate of material on the basis of the evidence obtained by inspecting only part of it. The evidence may be in the form of percentage of units nonconforming or may consist of measurements for each unit inspected. It is customary to denote the procedure, in the first case, as acceptance sampling by attributes and, in the second case, as acceptance sampling by variables.

Early in 1945 the Statistical Research Group, Columbia University, operating under the direction of the Applied Mathematics Panel, Office of Scientific Research and Development, was asked by the Navy to prepare a manual on sampling inspection, including tables, procedures, and principles. A manual on acceptance sampling by attributes was prepared first. The close of the war prevented completion of the work then under way on acceptance sampling by variables. The present volume is an expansion of the preliminary edition of the Navy manual.

The book has the subtitle: Principles, Procedures, and Tables for Single, Double, and Sequential Sampling in Acceptance Inspection and Quality Control Based on Percent Defective. The last four words should be emphasized since, in view of the main title, *Sampling Inspection*, the reader might otherwise expect to find some discussion of sampling inspection based on measurements. The reviewer's main criticism of this important book might be that no treatment of acceptance sampling by variables has been included. In view of the book's subtitle probably this is an unfair cavil and merely reflects the reviewer's disappointment that this companion problem still lacks the same comprehensive exposition which the present volume now makes available for acceptance sampling by attributes.

The three purposes of the book (as stated by the authors on p. 3) are: (1) "to explain the general principles of sampling inspection as a method for determining the quality and acceptability of material and as a method for controlling the quality of material"; (2) "to describe in detail a standard sampling-inspection procedure that may be readily applied to a wide variety of products"; (3) "to provide an extensive catalogue of sampling plans, usable both by those who adopt the standard procedure and by those who construct tailor-made procedures." With the restriction noted above, the book succeeds admirably in accomplishing all three purposes. It will meet the needs of three kinds of readers: those who want a ready-made sampling-inspection program, those interested in building their own programs, and students of acceptance sampling.

The book is divided into five parts. Part I consists of a chapter of introduction; Part II contains five chapters of discussion of the principles of sampling inspection for attributes; Part III, with seven chapters, gives a description of a standard sampling-inspection procedure; Part IV has four chapters on the construction of sampling tables and procedures. The foundation of the book, as noted in the preface, is the set of sampling tables presented in Part V.

As suggested in the Preface, Parts II, III, and IV have been written in such form that anyone needing a ready-made sampling inspection program can use Parts I, III, and V, omitting Parts II and IV; readers interested in building their own programs can substitute Part II for Part III; and, finally, those who want to see all of the technical details of tabular construction will find them in Part IV. Probably this plan of organization increases the scope of utilization of the book, but it requires that there be a considerable amount of repetition.

Ordinarily, the surest way to avoid pleasure from reading a novel is to read the last chapter first and certainly it is unorthodox to review a book in that order. This reviewer, however, could not resist the temptation to turn to Part V and, after much consideration, decided this would be an appropriate part to describe first.

There are four tables in Part V. Table 4, which is captioned "Sampling Plans and Operating-characteristic (OC) Curves, Classified by Sample-size Letter, and by Acceptable-quality Level (AQL) and Average Outgoing-quality Limit (AOQL)," consists of 159 pages. With a few exceptions, each page contains a single, a double, and a group sequential sampling plan together with the operating characteristic curves for each. The three plans are matched on the basis of their operating characteristic curves.

On page 288, a fairly representative page, there is a single sampling plan specifying a sample of 75 and an acceptance number of 1; a double sampling specifying a first sample of 50 and a second sample of 100, with acceptance number of 0 and rejection number of 3 for the first sample and with acceptance number of 2 and rejection number of 3 for the combined sample; a sequential plan, having seven steps of 20 units each, with decision assured at the end of the seventh sample. Printed in color, in the lower part of the page, are the operating-characteristic curves for the three plans.

The single and double plans, at least, are familiar since they appear in the Standard Sampling Inspection Tables issued by the Army Service Forces as part of ORD-M608–8. (This is not surprising since, in the words of the authors (p. vi) "the tables and procedures in this book are in large measure adapted from or based on Army Ordnance tables and procedures.") The double sampling plan, for example, is given in Table 1, Normal Lot-by-Lot Inspection, corresponding to an Acceptable Quality Level of .26 – 50% for major defects and sublots of size of 500 to 799. The approximate AOQL is recorded as 1%. In the volume under review, however, the classification is "Sample-size Letter: H; AQL class: 0.32–0.65; AOQL class: 0.90–1.5." The

classification by sample-size letter, then would seem to have replaced classification on the basis of lot size. Actually the change is not so direct, or so easy to dismiss.

In order to find the appropriate sampling plan in the set of Army Ordnance Tables, the type of sampling (single or double) and the acceptable quality level must be specified; the sublot size must be known; and information regarding the level and state of control of the manufacturing process should be available. For any particular type of sampling, sublot size and acceptable quality level, three plans are available; a "normal" plan, to be used when the process average is near the acceptable quality level; a "reduced inspection" plan, when the process is in control at a level significantly better than the acceptable quality level; and a "more exacting" plan, when the average quality of inspected product is significantly worse than the acceptable quality level. The "reduced inspection" plan, when appropriate, requires about one-fifth as much sampling as the "normal" plan. The "more stringent" plan uses the same sample sizes as the "normal" plan, but lower acceptance numbers, thus causing the rejection of a greater percentage of imperfect lots. The tables contain the criteria upon which a decision between plans is made for each lot which is submitted.

The first step in abstracting the appropriate sampling plan from the set of plans provided in the book under review might be to enter Table 1 of Part V. This table is captioned "Sample-size Letter, by Inspection Level and Size of Inspection Lot." There are five "inspection levels": I, II, III, IV, V. The amount of inspection at these levels is in the ratio $1:1\frac{1}{2}:2:3:4$ (approximately). A footnote states that "inspection level III will be used for most products for normal inspection," and, corresponding to this level and lot sizes 500–800 is found sample-size letter H, which appeared on page 288 in conjunction with "normal" Army Ordnance plans. Any of the other four levels, however, might be chosen as "normal," the choice depending on economic considerations.

In order to proceed further it is necessary to fix an Acceptable Quality Level. Table 2, "Summary of Sampling Plans Classified by Acceptable-quality Level and Sample Size Letter," has subtables for single-sampling, for double-sampling and for sequential-sampling. The specific plan to be adopted can be obtained here or direct reference may be made to Table 4, where three types of plans are matched as described above.

The reader who turns directly to the tables of Part V probably will be at a loss as to what inspection level to choose for "normal" inspection and, if he is familiar with other sets of sampling plans, he may wonder, perhaps, under what conditions there should be some change of plan in order to provide for "reduced" or "tightened" inspection. The term "inspection level" seems to be new. As explained on pages 45–46 and elsewhere, it is set for the product irrespective of the size of the lots into which the product is subdivided for inspection purposes. The higher the inspection level for a given lot size, the better will be the discrimination between good and bad lots. For

the same acceptable quality level and lot size the plans specified for the five inspection levels, provide, in general, for a decrease in risk of accepting unwanted product with an increase in inspection level. Inspection level is not to be confused with percentage of inspection; these tables follow Army Ordnance in providing for some increase in discrimination (and, therefore, sample size) as the lot sizes increase.

Footnotes in Table 1 direct the user to drop his inspection level two steps for reduced inspection, but to retain the normal level for tightened inspection. It is unfortunate that Table 1 contains no reference to Part III, Chapter 11 (pp. 120–125) for it is here that the authors tell when and how to change inspection plans. The criteria and procedures are substantially the same as given in the Ordnance Tables.

If rejection of lots is to be interpreted as sentencing lots to 100% inspection then it is important to know the Average Outgoing-quality Limit for material inspected. Table 3, "Sample-size Letters for Average Outgoing-quality Limit (AOQL) Classes and Acceptable-quality Level (AQL) Classes," supplies the required information. It may also be used to get sampling plans when an AOQL is specified.

To indicate further the difference of these tables from the Army Ordnance Tables it seems desirable to quote the authors (p. 142).

"The major respects in which the present collection of sampling plans differs from the Army Ordnance collection are as follows: (*a*) it includes sequential-sampling plans as well as single- and double-sampling plans; (*b*) it includes plans for smaller sample sizes—for example, the single-sampling in the main Army Ordnance single-sampling table (Table C) have sample sizes varying from 75 to 1,500; in Table 2 of this book, the sample sizes vary from 5 to 1,500; (*c*) it uses a different concept of AQL; (*d*) all the plans are contained in a single table (Table 2), whereas Army Ordnance uses supplementary tables for reduced and tightened inspection and for inspection for incidental defects."

It would be very culpable for the reviewer to imply that this book contains merely a set of sampling plans together with brief instructions regarding their utilization. On the contrary, the book treats the subject of sampling inspection by attributes in great detail. The bases for the decisions necessary in installing and operating acceptance sampling are discussed from almost every angle. The methods of constructing the sampling plans are carefully illustrated. Various possible alternatives are examined and the reasons why they were discarded are given in full. The reviewer has been unable to think of any question regarding the tables which has not been answered somewhere in the book. Nor has he entertained any criticism of choice of procedures without being able to discover that the authors have anticipated the possibility of just this criticism and have prepared a strong brief in support of their procedures. This insistence on complete coverage makes the book a valuable reference, especially in view of the scarcity of general publications in the field.

The introduction to the principles of sampling inspection for attributes which is given by Part II is excellent. Using well-chosen examples the authors describe single, double and group sequential sampling; introduce the reader to operating characteristic, average outgoing quality, and average amount of inspection curves; and discuss the various factors connected with the installation and operation of sampling plans. Methods of constructing control charts for fraction defective and for defects per unit are well presented and the use of sampling plans for control sampling is discussed. Finally there is a brief, but pointed appendix on the relation of sampling inspection to design specifications.

Part III describes how the set of tables in Part V can be used as the basis of an integrated inspection program. The considerations necessary in starting a sampling procedure are outlined and then the details of the operation of a sampling procedure are set down. Step-by-step the inspector is told what to do, how to do it, and why. Here he finds rules for drawing the sample, making his decision on the lot, calculating process average, changing from normal to reduced or tightened inspection, etc. This is the part of the book which activates the tables. It is an expert job and major excerpts will, no doubt, be quoted in many an inspection manual.

The somewhat limited class of readers who are concerned with methods of constructing sampling tables will welcome Part IV with enthusiasm. Up to present time detailed knowledge of the procedures used in calculating tables has been restricted largely to those intimately connected with the actual task. Given an individual sampling plan, it is relatively simple to write the formulas and make the calculations necessary to determine how the plan can be expected to operate; given rational requirements for an individual plan, it is not too difficult to write the equations which, if they could be solved, would lead to a plan exactly meeting the specifications; but, given alternative sets of desired properties, none of which can be satisfied exactly, the problem of discovering which properties can be imposed on a set of sampling plans at a satisfactory level of uniformity and then planning the execution of the necessary computations is neither small nor easy. Therefore, anyone who has this last problem to solve has the greatest interest in the philosophy, principles and techniques basic to previous solutions. The clarity and comprehensiveness of Part IV will be appreciated.

There are a few harmless misprints in the book, and there is an occasional statement with which the reviewer could not entirely agree. On page 12, for instance, the remark is made that "the acceptance or rejection of an entire inspection lot on the basis of the quality of a sample drawn from that inspection lot provides a strong incentive for the supplier to improve the quality of his product." If, as often happens, the customer's inspection is preceded by a final inspection by the producer, then it is the improvement of this final inspection for which the customer's rejections provide a strong incentive.

In an attempt at simplification the authors have decided to redefine ac-

ceptable-quality level (AQL) as "percentage of defective items in an inspection lot such that the sampling plan will result in the acceptance of 95 per cent of submitted inspection lots containing that percentage of defective items" (p. 144). This reviewer is not alone in deploring the placing of this additional, and seemingly unnecessary, hazard in the path of the lay reader of statistics.

But, in spite of his few minor criticisms, the reviewer is convinced that this is an outstanding book. Anyone who is at all concerned with the problem of acceptance sampling by attributes needs to be thoroughly conversant with its content.

REVIEW BY PAUL S. OLMSTEAD
Bell Telephone Laboratories, Inc., Murray Hill, New Jersey

THE authors state that *Sampling Inspection* presents "in a form suitable for the peacetime needs of industry, sampling procedures for use in either acceptance inspection or process control." Only two of the seventeen chapters deal with process control. Hence, it may be stated that, basically, the book is for a consumer who is unwilling to accept any evidence concerning the quality of a lot except that given by test of a sample of that lot. The plans associated with the procedures are devised to reject approximately 5% of lots at the Acceptable Quality Level (AQL—this term is defined in a new way) and to assure delivery of product in the long run, subject to additional work by the supplier, that will meet an Average Outgoing Quality Limit (AOQL). This incidentally places a limit on the Lot Tolerance Per Cent Defective (LTPD). Plans for reduced inspection on the basis of spaced continued satisfactory quality from one supplier aim to keep rejections at AQL and acceptances at LTPD at approximately the same levels as in normal plans. Plans for tightened inspection to protect against poor quality from a supplier with a previous record showing that his quality is normally poorer than AQL keep sample sizes the same but reduce in per cent the AQL, AOQL and LTPD. This places a very severe penalty on lots at AQL in many cases, reducing the chance of acceptance from 95% to as little as 10%. No consideration is given to the possible alternative of having suppliers submit evidence of control for consecutive lots produced prior to a current lot as partial evidence for the satisfactory quality of the lot submitted. Under this scheme, test by the consumer need be only a check inspection to verify that the lot submitted conforms to what might be expected on the basis of such evidence for control.

In addition to the sampling plans (including Operating-Characteristic (OC) Curves), which constitute almost half of the book, an account is given of methods of calculation and various bases and procedures for making use of the plans. It is pointed out that these plans relate only to the "quality of a lot" expressed by "the percentage of defective items it contains," where the quality of an individual item is expressed as "defective or nondefective." Part I is a short chapter of an introductory nature. Part II defines the major

concepts used in the book and discusses the principles of sampling inspection for attributes taken into account in the construction of the tables. It includes simple examples of single, double, and sequential sampling plans. Chapter 5, in particular, points out many of the difficulties associated with sampling inspection, dwelling at length on the vagaries of selecting a "random sample". The answer reached in one instance associated with spottiness is not apt to satisfy engineers who feel that one of the objects of sampling inspection is to discover evidence for just such a condition. Such evidence is often provided by keeping track of the order in which defectives are encountered. Order is not mentioned in this book. The use of sampling inspection data to obtain evidence for control is discussed briefly in Chapter 6. This chapter rightly points out that many engineers may not be satisfied with control charts based on data for which 5% of the points are expected outside control limits when the process is in fact under control. This is one consequence of the choice of definition of AQL. Incidentally, the authors use the words "take corrective action" as applying to this 5% of the points rather than the more appropriate "look for trouble." The engineer "takes corrective action" only when he finds that a criterion based on experience as associated with trouble is exceeded. In other cases, he may "look for trouble" but, ordinarily, would not find it necessary to "take corrective action." This part closes with a chapter dealing with the difference between design requirements and acceptance requirements.

Part III considers a "standard sampling inspection procedure." What is presented is a reasonably complete prospectus, with numerous examples, that may be followed as a model by those who are willing to accept the underlying principles mentioned above. Most engineers will find much of the material useful as a general guide to a larger number of things that may cause trouble in setting up inspection procedures. Part IV discusses the coverage of the tables and the methods used in computation, dealing particularly with shortcuts, limits of error, and approximations. Some mathematical statisticians may object to the use of the Poisson Exponential as an approximation to the Hypergeometric Series. Considerable space is devoted to showing how to compute the average amount of inspection (ASN—Average Sample Number) required to reach a decision to accept or reject for each of the three types of sampling. Readers should be careful not to confuse this average amount of inspection with the average amount of inspection required to give assurance that the quality of product meets the AOQL. The latter takes into account the detailing of rejected lots and therefore may differ but little for the three types of sampling. Comparative values of ASN are given for only a few of the sampling plans.

Part V is devoted to presentation of the tables. These, in most cases, show single, double and limited step curtailed sequential sampling plans for ranges in lot size (cross referenced to a sample size letter) and ranges in AQL and AOQL. On each page along with a table showing three sampling plans is a chart with the OC curves for these plans. These curves, being in color, are

subject to printing errors as explained on page x. The book also contains a consolidated list of references, a glossary, and an index. The relatively few errors, mainly associated with printing, should not cause the reader difficulty.

It is quite likely that many engineers will find the sampling tables useful in their work. The underlying philosophy associated with their use, however, may lead to a somewhat different "standard sampling inspection procedure" —one that would place more emphasis on process control and less on acceptance inspection than the one outlined in Part III of the book. Consequently, this reviewer would classify *Sampling Inspection* as a good reference book but not a textbook.

How to Chart: Facts From Figures With Graphs. *Walter E. Weld* (Investment Adviser, 227 Rawson Road, Brookline, Mass.). Preface by *A. C. Haskell* (Codex Book Co., Inc., Norwood, Mass.). Norwood, Mass.: Codex Book Co., Inc. (74 Broadway), 1947. Pp. xiv, 218. $3.00.

REVIEW BY KENNETH W. HAEMER
Statistician, American Telephone and Telegraph Company
195 Broadway, New York 7, New York

THIS is a book for the businessman who knows only a little about charts or about how to make them. On the other hand, several of the seventeen chapters discuss chart uses and methods that even the experienced chartist—businessman or not—might employ with profit.

Item number one on the credit side is the author's style of writing. It is easy to read, free of unnecessary technical terms, and effective in conveying the author's ideas. It can be summarized as conversational rather than academic.

The illustrations are excellent, plentiful, and well drawn; they accurately illustrate the point under discussion, a feature more often absent than present in books about charts. It is a sad paradox that so many books on graphics are so poorly illustrated: Mr. Weld is to be congratulated on refusing to follow this "do as I say—don't do as I do" line of least resistance.

Throughout the book, the charts are explained in terms of specific business applications. Although this approach has the mild disadvantage of seeming to restrict the chart to only one type of use or application, it does succeed in bringing the charts to life and in pointing up exactly how each fits a particular type of business activity. It should be an effective means of introducing charts to any businessman who is just beginning to discover the extent of their usefulness.

The principles of chart presentation are skillfully explained. It is perhaps regrettable that Mr. Weld does not accept the principle that simple bar comparisons (1-scale measurement) are generally best shown as horizontal bars, and that the vertical arrangement is best reserved for 2-scale bar presentation. Also, in his discussion of multiple-vertical-scale presentation he

has mentioned only the advantages and has neglected to point out the considerable disadvantages. In almost all other respects, however, the graphic principles set forth are beyond reproach.

Considerable space is given to instruction on how to prepare charts. Although these instructions naturally reflect the author's own experience, and occasionally ignore the fact that there are alternatives, they should be extremely helpful to the beginning chart maker. Some of these practical hints are especially valuable because they so seldom find their way into print. For example, Mr. Weld points out that since the color light blue does not photograph, lines and other markings that are *not* to appear on the reproduced copies should be made on the original with a light blue pencil. This may seem to be a small point, but it is one that saves troubles and dollars, and one that incidentally demonstrates the author's practical point of view.

This book focuses attention on three types of charts that in many other books are discussed only briefly or are not mentioned at all. Each of them is a type developed from a special kind of scale: of the three, the semi-logarithmic chart properly receives the most emphasis. However, the other two, probability scale and reciprocal scale charts, are discussed at sufficient length to explain how and when they can be used to advantage.

The first of the two chapters on the semi-logarithmic chart provides an excellent introduction to the mechanics of this form of presentation. Of course, it is possible to quibble with some of the things the author has to say about the semi-logarithmic chart: it is a subject that can be quibbled about with very little effort. For example, he says " . . . it is often desirable and extremely helpful to know the *rate* of change, or percentage of gain or loss. The kind of chart which gives this information is called a ratio (or semi-logarithmic) chart." This reviewer holds to the belief that a semi-log chart shows only *relative* rate of change. But the preciseness of the term is indeed a minor matter and does not detract from the general soundness of the exposition. In Chapter 7, which discusses the uses of semi-log charts, Mr. Weld not only explains when the type of presentation is appropriate, but repeatedly points out that it should not be used with the mistaken idea that it can simultaneously serve the purpose of an arithmetic chart. This one point alone, if thoroughly absorbed by the reader, would be worth the price of the entire book.

The author's discussion of the probability scale and its uses in business charting (Chap. 8) is easy to read and is informative. Here, as elsewhere throughout the book, the statistical propriety of the applications cited, or the conclusions concerning these uses, are not considered to be within the scope of this review: Judged solely as an explanation of its function as a tool of graphic analysis, this chapter is a good introduction to this type of chart. Although the reciprocal chart is primarily a means of computation rather than of presentation, Mr. Weld makes a good case for devoting a chapter to it (Chap. 9). His explanation of its wide usefulness should encourage a more extensive use of this little-known chart.

In this reviewer's opinion, two other chapters deserve special mention. Chapter 15, "Charts of Various Kinds," is a quick review of assorted chart varieties—from progress charts to trilinear charts—which the author feels warrant inclusion. This chapter provides a useful, if limited, supplementary "idea" file. Chapter 16, "Exhibits Good and Bad," is a chart clinic in which seventeen selected examples are subjected to critical examination. These should be two of the most valuable chapters in the book for the novice in chart presentation.

Since it is inevitable that author and reviewer seldom agree on all things, some chips must fall on the unfavorable side.

To this reviewer, the emphasis in the book seems faulty. It appears to be designed to focus attention on the major interests of the author rather than on the probable interests of the reader. Because of the author's very special field of interest, these two are unlikely to be the same.

For example, stock market analysis charting and business forecasting fill three chapters and parts of others. This is two chapters more than is allowed to any other subject except semi-logarithmic charts, and seems out of proportion in an introductory discussion to business charting. Indeed, this ardent exposition seems appropriate for only one group of businessmen —investment analysts—among which number the author presumably is one. Although other businessmen readers will no doubt be intrigued to learn that there are stock market patterns known as "The Head & Shoulders," and "The Rounding Bottom," to mention two of many, few are likely to feel an overwhelming desire to put this special knowledge into use.

On the other hand, the subject of map presentation is covered with one illustration and a few sentences. So fragmentary a treatment of so important a subject seems scarcely worth including.

In several spots, the sequence seems unnatural. Stock market charting (Chaps. 10–11) and Business Forecasting (Chap. 13), are separated by a chapter on "What Should Be Charted to Help the Business Manager." Chapter 14, "Making and Reproducing Graphs," is sandwiched between "Business Forecasting" and "Charts of Various Kinds."

There is a considerable amount of overlapping. This need not be a fault if it is carefully planned with a legitimate purpose, but in this case much of it has the appearance of being partly incidental and partly accidental. For example, one of the subheads under Chapter 17 is "Procuring Suitable Chart Sheets." Under Chapter 14 there is a subhead "Selection of Suitable Chart Paper." In fact, the Codex Company or its products are mentioned one or more times on 18 pages throughout the book. It is only fair to point out that several of these mentions include invitations to call on the Codex Company for further advice or guidance. Nevertheless, this reader found it disturbing to have the publisher's name thrust forward with such artless frequency. Another example of overlapping: One of the items covered in Chapter 4 is the matter of choosing lines or spaces to represent units of time. Although no mention is made in the chapter heading, Chapter 5 also discusses

this problem at some length. There is, of course, some advantage in repetition at strategic points, but this reader for one prefers to find most of the information on a given subject set down on adjacent pages.

Although it makes for ease of reading and a fast moving exposition, the author's penchant for singling out one device or one method and neglecting to mention the alternatives, does a disservice to the reader who wants to know the best way rather than just a good way. For example, when discussing lettering guides, Mr. Weld mentions only Wrico lettering guides. In spite of the fact that this reviewer leans toward Wrico guides, it would seem desirable to tell the reader at least what others exist, and in order to be of maximum help to explain the advantages and disadvantages of each.

This reviewer is of the opinion that no one book about charts can hope to answer all chart questions. The design and use of charts involves so many decisions based on judgment and on knowledge outside the chart itself, that any book, however extensive, must leave many questions unanswered. The author of a book on charts is confronted with three major choices: a thorough job on principles, applying to all fields of charting; a thorough job on the detailed procedures for a limited field of charting; or a compromise combining some of each. In his present book, Mr. Weld has chosen to follow the latter course. He has succeeded very well in presenting the principles of simple charting, spiced with enough specific applications to make them appetizing and easy to digest. Although it is directed primarily to businessmen, this book has something to say also to nonbusiness chart makers.

In addition to providing a helpful introduction to the beginner, this book fits neatly onto the bookshelf, next to more voluminous works, where it should serve as a compact reference to the author's special field of interest.

NOTE: As a result of the reduction in the number of pages in this issue of the *Journal*, the Statistical Methodology Index is being omitted. Readers are invited to write to the Secretary's office concerning the desirability of continuing this series of bibliographies of the literature on statistical methodology.

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION

CONTENTS OF VOLUME 43

THE 107TH ANNUAL MEETING

PROGRAM. 299

MINUTES OF THE ANNUAL BUSINESS MEETING 308

ARTICLES 1, 211, 357, 497

BOOK REVIEWS 165, 312, 461, 635

LETTERS ABOUT BOOKS 172

STATISTICAL METHODOLOGY INDEX 174, 356, 500, 668

INDEX TO VOLUME 43, 1948

ARTICLES, BY AUTHOR 620

BOOK REVIEWS, BY AUTHOR 621

REPORTS AND OFFICIAL NOTICES

REPORT OF THE BOARD OF DIRECTORS 311

REPORT OF THE SECRETARY 314

REPORT OF THE TREASURER 314

REPORT OF THE AUDITORS 315

COMMITTEES AND REPRESENTATIVES FOR 1947 307

INDEX TO VOLUME 43: 1948

ARTICLES

ARNOLD, DAVID R., and SPURR, WILLIAM A., *A Short-Cut Method of Fitting a Logistic Curve* 127

BOWKER, ALBERT H., *A Test for Symmetry in Contingency Tables* . . . 572

*BROWN, WEIR M., *Measuring Physical Inventories* 377

*CAMPBELL, GEORGE C., *Problems with Sampling Procedures for Reserve Valuations* 413

*CODY, DONALD D., *Sampling Errors in Mortality and Other Statistics in Life Insurance* 442

DEEMER, WALTER L., JR., *The Use of Mark Sensing in a Large Scale Testing Program* 40

DIXON, W. J., and MOOD, A. M., *A Method for Obtaining and Analyzing Sensitivity Data* 109

*EPSTEIN, BENJAMIN, *Application of the Theory of Extreme Values in Fracture Problems* 403

FINNEY, D. J., *Main Effects and Interactions* 566

*GREVILLE, THOMAS N. E., *Recent Developments in Graduation and Interpolation* 428

GRUBBS, FRANK E., *On Estimating Precision of Measuring Instruments and Product Variability* 243

Correction 564

HARRIS, MARILYN; HORVITZ, D. G., and MOOD, A. M., *On the Determination of Sample Sizes in Designing Experiments* 391

HOCHSTIM, JOSEPH R., and STOCK, J. STEVENS, *Commercial Uses of Sampling* 509

HORVITZ, D. G.; HARRIS, MARILYN, and MOOD, A. M., *On the Determination of Sample Sizes in Designing Experiments* 391

KIRCHEN, C. J., *Levels of Significance for Variance Ratio of Two Samples of Equal Size* 564

LASKEN, HERMAN, *Direct Determination of Compass Settings for Proportionate Area Pie-Charts* . . . . . JOHN V. SPIELMANS 94

LEBERGOTT, STANLEY, *Earnings of Nonfarm Employees in the U.S., 1890–1946* 74

*LEW, EDWARD A., *Mortality Statistics for Life Insurance Underwriting* . . 274

MOOD, A. M., and DIXON, W. J., *A Method for Obtaining and Analyzing Sensitivity Data* 109

*MCDIARMID, FERGUS J., *Life Company Investments and the Capital Markets* 265

*MILLER, MORTON D., *Statistical Problems in Developing Voluntary Medical Expense Insurance Plans* 290

MOOD, A. M.; HARRIS, MARILYN, and HORVITZ, D. G., *On the Determination of Sample Sizes in Designing Experiments* 391

MOSTELLER, FREDERICK, *On Pooling Data* 231

NIESSEN, ABRAHAM M., *Actuarial Estimates for Public Sickness Insurance Plans* 61

NOVICK, DAVID, and STEINER, GEORGE A., *The War Production Board's Statistical Reporting Experience* 201, 463, 575

*PARADISO, LOUIS J., *Significance of Inventories in the Current Economic Situation* 361
RICHARDSON, LEWIS F., *Variation of the Frequency of Fatal Quarrels with Magnitude* 523
SCHUYLER, GARRET L. *The Ordering of n Items Assigned to k Rank Categories by Votes of m Individuals* 559
*SNEDECOR, GEORGE W., *A Proposed Basic Course in Statistics* . . . 53
SPIELMANS, JOHN V., *Profile Graphs* 96
SPURR, WILLIAM A., and ARNOLD, DAVID R., *A Short-Cut Method of Fitting a Logistic Curve* 127
STEPHAN, FREDERICK F., *History of the Uses of Modern Sampling Procedures* 12
STOCK, J. STEVENS, and HOCHSTIM, JOSEPH R., *Commercial Uses of Sampling* 509
*THORP, WILLARD L., *Statistics and Foreign Policy* 1
THURSTONE, L. L., *The Edge-Marking Method of Analyzing Data* . . . 451
WARBURTON, CLARK, *Bank Reserves and Business Fluctuations* . . . 547

BOOK REVIEWS

ABAY, GYULA, *Bevezetés a Statisztika Tudományába: Rész I* . EUGENE LUKACS 135
ALBERT, ADRIAN A., *College Algebra* . . . . . PAUL S. DWYER 136
American Society for Quality Control, First Annual Convention and Second Midwest Quality Control Conference, June 5–6, 1947, Conference Papers . . . . . H. A. FREEMAN 353
. . . . . EUGENE L. GRANT 354
BURTON, CLEMENT, *A Concise Manual of Statistics: With Special Reference to the Requirements of Students for Municipal Examination, Second Edition* . . . . . HARRY PELLE HARTKEMEIER 137
CHAPIN, F. STUART, *Experimental Designs in Sociological Research* . . .
. . . . . O. KEMPTHORNE 489
COLUMBIA UNIVERSITY STATISTICAL RESEARCH GROUP, Applied Mathematics Panel, Office of Scientific Research and Development, Edited by H. A. FREEMAN, *Sampling Inspection: Principles, Procedures and Tables for Single, Double and Sequential Sampling in Acceptance Inspection and Quality Control Based on Percent Defective* . . .
. . . . . EDWIN G. OLDS 605
. . . . . PAUL S. OLMSTEAD 611
DAHLBERG, GUNNAR, *Mathematische Erblichkeitsanalyse von Populationen* .
. . . . . OYSTEIN ORE 139
DAVIES, OWEN L., editor. *Statistical Methods in Research and Production: With Special Reference to the Chemical Industry* . . . . . .
. . . . . D. B. DELURY and A. E. R. WESTMAN 318
. . . . . E. L. LEHMANN 320
DEWEY, EDWARD R., and DAKIN, EDWIN F., *Cycles: The Science of Prediction* . . . . . MILTON FRIEDMAN 139
. . . . . MAX SASULY 141
EISENHART, CHURCHILL; HASTAY, MILLARD W., AND WALLIS, W. ALLEN, editors. Statistical Research Group, Columbia University, *Selected Techniques of Statistical Analysis: For Scientific and Industrial Research*

and Production and Management Engineering . . HENRY SCHEFFÉ 321
FINNEY, D. J., *Probit Analysis: A Statistical Treatment of the Sigmoid Response Curve* . . . . . MARGARET MERRELL 145
. . . . . JOSEPH BERKSON 148
FISHER, RONALD A., *The Design of Experiments, Fourth Edition* . . .
. . . . . A. E. BRANDT 323
FLORENCE, P. SARGANT, *Investment, Location, and Size of Plant: A Realistic Inquiry into the Structure of British and American Industries* . . .
. . . . . EDGAR M. HOOVER 601
FURLAN, L. V., *Das Harmoniegesetz der Statistik: Eine Untersuchung über die metrische Interdependenz der sozialen Erscheinungen* . . . . .
. . . . . HILDA GEIRINGER 325
GUERREIRO, AMARO D., *Manuel de Estatistica* . . . HUGO MUENCH 153
HASTAY, MILLARD W.; EISENHART, CHURCHILL, and WALLIS, W. ALLEN, editors. Statistical Research Group, Columbia University, *Selected Techniques of Statistical Analysis for Scientific and Industrial Research and Production and Management Engineering* . . HENRY SCHEFFÉ 321
HAYS, SAMUEL, *An Outline of Statistics, Third Edition* HAROLD NISSELSON 155
. . . . . EDGAR Z. PALMER 156
HEIDINGSFIELD, MYRON S., and BLANKENSHIP, ALBERT B., *Market and Marketing Analysis* . . . . . LESTER R. FRANKEL 157
. . . . . PHILIP J. MCCARTHY 160
HILL, A. BRADFORD, *Principles of Medical Statistics. Fourth Edition* . .
. . . . . J. YERUSHLMY 492
INSTITUTE OF MATHEMATICAL STATISTICS COMMITTEE ON THE TEACHING OF STATISTICS, Report of the, *The Teaching of Statistics* . . . .
. . . . . .TRUMAN L. KELLEY 493
. . . . . ALAN E. TRELOAR 597
KELLEY, TRUMAN LEE, *Fundamentals of Statistics* . . . CYRIL BURT 328
. . . . . HORACE W. NORTON 330
KENNEY, JOHN F., *Mathematics of Statistics*, Part One, Second Edition . .
. . . . . HOWARD LEVENE 332
. . . . . RONALD W. SHEPHARD 334
LEVENS, A. S., *Nomography* . . . . . M. KAC 496
MORTARA, GIORGIO, *O Custo de Produção do Homem Adulto e Sua Variação em Relação à Mortalidade* . . . . . T. N. E. GREVILLE 161
PEATMAN, JOHN GRAY, *Descriptive and Sampling Statistics* . . . . .
. . . . . LEON FESTINGER 335
. . . . . FREDERICK F. STEPHAN 337
ROYAL STATISTICAL SOCIETY'S COMMITTEE ON THE TEACHING OF STATISTICS, *The Teaching of Statistics in Universities and Colleges* TRUMAN L. KELLEY 493
. . . . . ALAN E. TRELOAR 597
SAMUELSON, PAUL ANTHONY, *Foundations of Economics Analysis* . . .
. . . . . GEORGE J. STIGLER 603
. . . . . GERHARD TINTNER 497
SCARBOROUGH, J. B., and WAGNER, R. W. *Fundamentals of Statistics* . .
. . . . . LEO AROIAN 339
SCHWENG, LORAND, *Statisztika Módszertani Alapvetés* . EUGENE LUKACS 135
SMART, L. EDWIN, and ARNOLD SAM, *Practical Rules for Graphic Presentation of Business Statistics* . . . . . RUDOLF MODLEY 341

SMITH, EDWARD S., *Control Charts: An Introduction to Statistical Quality Control* . . . . . JOSEPH M. JURAN 343
. . . . . MASON E. WESCOTT 343
THOMPSON G. CLARK, *Forecasting Sales* . . . . CHARLES F. ROOS 163
. . . . . H. D. WOLFE 165
THIONET, PIERRE, *Méthodes Statistiques Modernes des Administrations Fédérales aux États-Unis* . . . . . W. EDWARDS DEMING 346
WELD, WALTER E., *How to Chart: Facts From Figures With Graphs* . .
. . . . . KENNETH W. HAEMER 613
WESTMAN, A. E. R., *An Extension Course in Statistical Quality Control* .
. . . . . A. C. COHEN, JR. 350
. . . . . PAUL PEACH 351
WRIGHT, WILSON, *Forecasting for Profit: A Technique for Business Management* . . . . . WILLIAM A. SPURR 167
. . . . . VICTOR VON SZELISKI 168
ZETTERLUND, STIG, *Toleranser och Passninger*. . . . CECIL C. CRAIG 352

GEORGE BANTA PUBLISHING COMPANY, MENASHA, WISCONSIN